

EXAMINING THE NATURE OF VARIABILITY IN GENDER AND NUMBER  
AGREEMENT IN NATIVE AND NON-NATIVE SPANISH

BY

Beatriz López Prego

Submitted to the graduate degree program in Linguistics  
and the Graduate Faculty of the University of Kansas in  
partial fulfillment of the requirements for the degree of  
Master's of Arts

Committee:

---

Chair: Alison Gabriele

---

Member: Robert Fiorentino

---

Member: Sara Rosen

Date defended: 02/08/12

The Thesis Committee for Beatriz López Prego certifies  
that this is the approved version of the following thesis:

EXAMINING THE NATURE OF VARIABILITY IN GENDER AND NUMBER  
AGREEMENT IN NATIVE AND NON-NATIVE SPANISH

Committee:

---

Chair: Alison Gabriele

---

Member: Robert Fiorentino

---

Member: Sara Rosen

Date approved: 02/08/12

## ABSTRACT

We explored the nature of morphological variability in English-speaking learners of Spanish, examining whether the variability is systematic and whether performance is impacted by task demands. McCarthy (2008) proposed that ‘default’ errors, referring to the incorrect extension of masculine/singular morphemes to contexts requiring feminine/plural agreement, are more frequent than the opposite pattern, called ‘feature clash’ errors. According to representational accounts of morphological variability, the source of the errors in L2 learners is the inability to acquire features absent from the L1. In contrast, computational accounts argue that this variability in learners is due to processing issues. In order to test these hypotheses, our L2 learners took either a Speeded Grammaticality Judgement Task (GJT) or an Untimed GJT. Three groups of Spanish natives took a Speeded GJT at three different rates of presentation (Speeds 1, 2, and 3) in order to examine whether natives under pressure perform similarly to learners (Hopp, 2010). The results show quantitative effects of speed for both natives and learners with decreased rates of accuracy in both groups. Natives were significantly less sensitive to errors as speed increased. For the learners, all proficiency levels showed increased sensitivity in the untimed task, with individuals in the advanced group showing target-like performance. With respect to the qualitative nature of the errors, differences between ‘default’ and ‘feature clash’ errors emerged for the natives in Speed 2 and the intermediate L2 group in the Speed GJT. The results for both groups showed more sensitivity to feature-clash than default errors for number. For gender, L2 learners showed the opposite pattern, while no differences were observed in native speakers. Overall, our results are more consistent with computational accounts of morphological variability.

## ACKNOWLEDGMENTS

I would like to thank the Linguistics Department at the University of Kansas for making this project possible. I am very grateful to my adviser, Alison Gabriele, for her dedication, her patience, and her guidance every step of the way. Many thanks to my committee members, Robert Fiorentino and Sara Rosen, for their advice and excellent feedback.

I would also like to thank my fellow graduate students, José Alemán Bañón, for his generous contribution to the stimuli; Stephen Politzer-Ahles, Alonso Canales Viquez, and Kristi Bond, for their comments and ideas on stimuli design and statistical analyses; and the Research in Acquisition and Processing seminar group for their feedback and encouragement.

For their invaluable help with recruiting L2 learners and for their participation as subjects in the study, I am very thankful to Amy Rossomondo and the GTAs of the KU Department of Spanish and Portuguese. I also owe thanks to Jonathan Perkins and the staff from the EGARC labs at KU, for their assistance with data collection. Thank you guys for making things so much easier and faster.

Special thanks to Carlos Acuña Fariña, who was instrumental in recruiting the native speakers as well as in providing the equipment for the experimental sessions, even before my arrival in Spain. Without your help this project wouldn't be a reality. I am very grateful to Natalia Rega González, who was also very helpful with recruiting participants, and assisted with every aspect of data collection in Spain. Thanks for being such a good friend and for making this experience so fun.

Finally, I would like to thank my family and friends for all their support, encouragement, and for giving me the strength I sometimes couldn't find within myself. Covadonga López

Prego, Javier Busto Mata, Laura Alonso Prieto, Alba Carro Pérez, Henrique Ferreiro García, Alejandra Fontán, Lucía Gómez Yáñez, Miriam Lago Iglesias, Marta Legaspi García, Xiana Melón Rodríguez, Mila G. Moa Cagiao, Ana Mosquera Leal, Alex Parente, Pilar De Valcárcel Sánchez, Gloria Prada Hervella, Ale Riveiro Míquez, Mónica Rodríguez Fernández, Miriam Vázquez Rodríguez, thanks for believing in me.

## DEDICATION

To my parents Ramón and María José.

Gracias por inculcarme la disciplina que tan valiosa ha sido para escribir esta tesis;

por vuestros sacrificios para que pudiera optar a todo lo que la vida ofrece;

por animarme siempre a ser lo mejor que puedo ser.

## TABLE OF CONTENTS

1. Introduction.....	1
2. Gender and Number in Spanish .....	4
3. Agreement in Spanish .....	6
3.1 Number Agreement.....	6
3.2 Gender Agreement.....	7
4. Approaches to Morphological Variability .....	9
4.1 Representational Accounts.....	10
4.2 Computational Accounts.....	13
5. Variability on Agreement in L2 Spanish .....	17
5.1 Distributed Morphology.....	18
5.2 McCarthy (2008).....	22
6. Variability in Native Speakers.....	26
7. The Present Study .....	32
7.1 Research Questions and Hypotheses .....	32
7.1.1 L2 Learners .....	32
7.1.2 Natives .....	34
7.2 Method .....	35
7.2.1 Stimuli.....	35
7.2.2 Tasks .....	39
7.2.2.1 Speeded GJT .....	40
7.2.2.2 Untimed GJT.....	42

7.2.3	Participants.....	42
7.2.3.1	L2 Learners .....	42
7.2.3.2	Native Speakers .....	44
7.2.4	Procedures.....	45
8.	Analysis and results .....	46
8.1	L2 Learners .....	48
8.1.1	Speeded GJT .....	48
8.1.1.1.	Overall Results.....	49
8.1.1.2	Group Results.....	50
8.1.2	Untimed GJT.....	52
8.2	Native Speakers .....	53
9.	Discussion.....	56
9.1	L2 Learners .....	57
9.2	Native Speakers .....	61
10.	Conclusion .....	63
	References.....	64
	APPENDIX. Target stimuli .....	70



## 1. Introduction

As has been widely reported in the literature on second language (L2) acquisition, morphology is one of the most problematic aspects of a second language to master (Franceschina, 2002; Hawkins & Chan, 1997; Hopp, 2010; Keating, 2009; Lardiere, 2000; McCarthy, 2007; McDonald, 2006; Montrul et al., 2008; Prévost & White, 2000; White et al., 2004, etc.). The complexity of the matter lies in part on the need for agreement between words, that is, in the dependency of the morphological form of a word (the *target* of agreement, determiners and adjectives among others) on another word (the *trigger* of agreement, usually nouns). Depending on whether this dependency is correctly realized or not, the same learner will show correct and incorrect instances of agreement in the same context, even when the same grammatical feature is involved.

To illustrate, in the first sentence of example (1) below the learner correctly establishes the dependency between the third person singular subject and the form of the verb, which shows the third person singular marker *-s*. However, in the second sentence, the learner fails to establish the Subject-Verb (S-V) dependency again, thus showing incorrect agreement. Similarly, in the first part of example (2), from Spanish, both the determiner and the adjective show the *-a* marker corresponding to agreement with the feminine, singular noun *falda*. Nevertheless, in the second part of the example, where the learner uses a pronoun to refer to the noun, she uses the masculine, singular form *lo*, reflecting a gender mismatch with the noun.

(1) My cousin **runs** every morning. He usually **\*run\_** approximately a mile.

(2) Tengo una falda roja. \*Me lo pongo en verano.  
I have a<sub>FemSg</sub> skirt<sub>FemSg</sub> red<sub>FemSg</sub> \*On me it<sub>MSg</sub> I put in summer.  
'I have a red skirt. I wear it in summer.'

The variability illustrated in the examples above extends to all types of inflection including case, gender, number, or person among others, with learners inconsistently supplying agreement morphology in targets such as determiners, adjectives, verbs, or pronouns.

The fact that this phenomenon has been attested in advanced learners in both production and comprehension (Franceschina, 2002; McCarthy, 2008) makes two questions especially prominent: is it possible to acquire morphology in the L2 to native-like levels? And what is the source of agreement errors in learners?

In response to these questions, some researchers adopt a **representational** approach (Hawkins & Chan, 1997; Hawkins, 2001; Franceschina, 2002; Tsimpli and Dimitrakopoulou, 2007), which proposes that features not instantiated in the native language (L1) are not acquirable in a second language after a critical period. That is, these theories posit a deficit in the abstract representation of L2 features due to late acquisition as the source of morphological variability. In this case, errors are expected to appear even in advanced learners if the L1 and L2 do not share the same features, and if acquisition started after a proposed critical period.

Other researchers adopt a different approach, and argue for a **computational** explanation of morphological variability (Hopp, 2010; Keating, 2009; Lardiere, 2000; McDonald, 2006; Prévost & White, 2000; White et al., 2004). Contrary to representational accounts, computational accounts attribute the source of the observed variability to performance or processing issues. These issues may come up when learners try to either access or map the appropriate features onto vocabulary items. Thus, these accounts propose that learners' inconsistency in supplying agreement morphology does not come from representational issues, but from computational difficulties.

Another important factor to be considered in this topic is the nature of learner variability. Many have noticed that agreement errors in Spanish tend to occur in a specific direction, that is, they appear to be systematic (McCarthy, 2008; White et al., 2004; Montrul et al., 2008). Error types like (3), where the masculine and singular forms are overextended to feminine, and plural contexts, seem to occur more frequently than (4), where feminine and plural forms are incorrectly used in masculine, and singular contexts:

- (3) La casa es \*antiguo, pero tiene unas vistas \*bonita.  
 The house<sub>FemSg</sub> is old<sub>MSg</sub> but it has some views<sub>FemPl</sub> beautiful<sub>FemSg</sub>  
 ‘The house is old, but it has some beautiful views.’
- (4) El libro es \*novedosa y tiene un argumento \*ingeniosos.  
 The book<sub>MSg</sub> is innovative<sub>FemSg</sub> and it has a plot<sub>MSg</sub> witty<sub>MPl</sub>  
 ‘The book is innovative and it has a witty plot.’

These patterns are argued to reflect use of default morphology by learners. Interestingly, recent studies have shown that native speakers exhibit similar patterns of variability *when under processing pressure* (Hopp, 2010; McDonald, 2006). In the present study we explore use of default morphology in learners and native speakers of Spanish in relation to different task demands, in order to shed some light on the debate regarding the sources of morphological variability and the two main explanations proposed. Importantly, in our study we test learners whose L1 is English, a language that instantiates number, but not gender, in contrast to Spanish, their L2, which instantiates both gender and number.

This thesis is organized into the following sections: Section 2 outlines the gender and number systems in Spanish; Section 3 discusses how gender and number agreement is established in Spanish; Section 4 introduces some of the L2 theories and studies that have attempted to explain agreement variability in L2 learners; Section 5 explains the specific account of learner variability that we test in our study; Section 6 motivates our methodology by

introducing two studies that have examined morphological variability in native speakers; Section 7 describes the present study; Section 8 presents the statistical analyses and results; Section 9 discusses our results and how they fit in the literature; Section 10 summarizes the main findings and conclusions of the study.

## 2. Gender and Number in Spanish

There are two genders in Spanish: masculine and feminine. Canonically, masculine nouns end in *-o* (*el niño*, *the boy*), and feminine nouns end in *-a* (*la niña*, ‘the girl’). There are many exceptions to this rule, with nouns ending in *-e* (*la noche*, ‘the night’<sub>fem</sub>), consonant or zero marker (*la tesis*, ‘the thesis’), or even the ending that would normally correspond to the opposite gender according to the canonical rule (*la moto*, ‘the motorcycle’<sub>fem</sub>; *el día*, ‘the day’<sub>masc.</sub>). According to Harris (1991) these endings are not gender markers, but morphemes, and therefore, they can occur in any word, although there is a tendency for masculine nouns to end in *-o*, and feminine nouns to end in *-a*. The present study focuses on canonical nouns only, in order to ensure that we are testing gender agreement, and not gender assignment knowledge of these special items.

Another irregularity worth mentioning is the masculine determiner, which takes the form *el* in the singular (example 5), and therefore, does not show the usual *-o* ending, as opposed to its plural form *los* (example 6). The feminine determiner, on the other hand, is always regular, and thus shows the *-a* feminine marker both in the singular (example 7) and the plural (example 8). Some examples with the definite articles and canonical nouns are provided below:

- (5) El suelo es rojoo.  
The<sub>MSg</sub> floor<sub>MSg</sub> is red<sub>MSg</sub>.
- (6) Los suelos son rojos.  
The<sub>MPI</sub> floors<sub>MPI</sub> are red<sub>MPI</sub>.
- (7) La casa es rojaa.  
The<sub>FemSg</sub> house<sub>FemSg</sub> is red<sub>FemSg</sub>.
- (8) Las casas son rojas.  
The<sub>FemPl</sub> houses<sub>FemPl</sub> are red<sub>FemPl</sub>.

We find a similar irregularity in the indefinite articles *un*, *una*, *unos*, *unas*, where, again the masculine, singular form lacks the canonical *-o* ending (example 9). As with the definite articles, the rest of the paradigm is regular. Some examples are provided below:

- (9) un libroo  
a<sub>MSg</sub> book<sub>MSg</sub>
- (10) unos libross  
some<sub>MSg</sub> books<sub>MSg</sub>
- (11) una libretaa  
a<sub>FemSg</sub> notebook<sub>FemSg</sub>
- (12) unas libretass  
some<sub>FemPl</sub> notebooks<sub>FemPl</sub>

With regards to number, the Spanish system has singular and plural. The canonical singular form is unmarked, while the canonical plural is formed by adding an *-s* to the root of the noun (examples 6, 8, 10, and 12 above). There are three exceptions to this rule: nouns ending in a consonant other than *-s* form their plural in *-es* (*el camión*, ‘the truck’- *los camiones*, ‘the trucks’); nouns ending in *-s* have the same form in singular and plural (*la tesis*, ‘the thesis’-*las tesis*, the theses’); multisyllabic nouns ending in a stressed vowel take either *-s* or *-es* (*el bisturí*, ‘the scalpel’-*los bisturís/bisturíes*, ‘the scalpels’). In the present study we only included nouns

with canonical plurals, as in examples (6), (8), (10), and (12) above, for the same reason for which we only included nouns with canonical gender, namely, testing agreement exclusively.

### 3. Agreement in Spanish

In Spanish, pronouns, determiners, and adjectives must agree in gender and number with the noun triggering agreement. Agreement takes place both within and across phrases. Example (13) below shows agreement within a Determiner Phrase (DP). Both the determiner and the adjective show the same gender and number as the feminine, singular noun *casa*.

- (13) la            casa            roja.  
       the<sub>FemSg</sub> house<sub>FemSg</sub> red<sub>FemSg</sub>  
       ‘the red house’

Example (14) shows agreement both within, and across phrases. The determiner is the head of the DP, the same phrase where the noun is a complement (within-phrase agreement), whereas the adjective is a predicative complement within the Verb Phrase (VP) (across-phrase agreement). Both the determiner and the adjective agree with the masculine, plural noun *coches*.

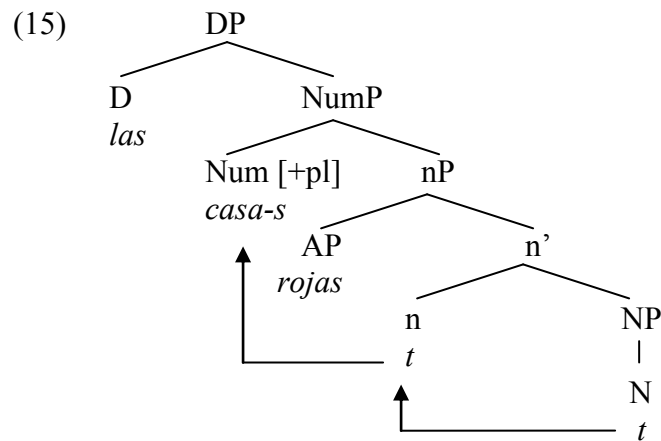
- (14) Los        coches son rojos.  
       The<sub>MPi</sub> cars<sub>MPi</sub> are red<sub>MPi</sub>  
       ‘The cars are red.’

In the present study we will analyze across-phrase agreement on predicative adjectives, similar to example (11). The rules for gender and number marking on adjectives are the same as the rules for nouns described above.

#### 3.1 Number Agreement

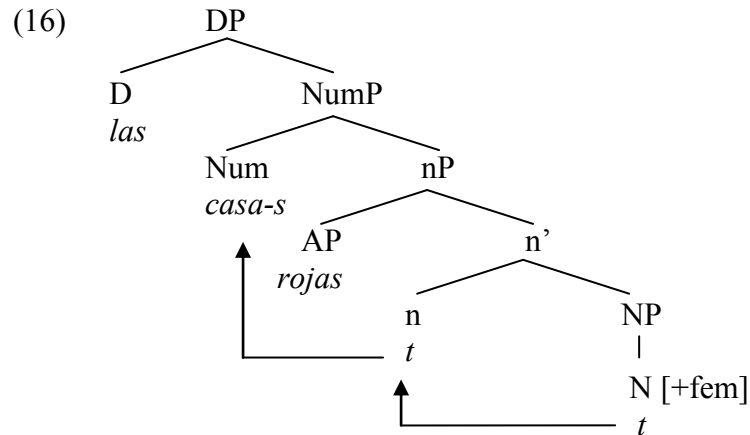
Following Chomsky’s (1995) theory of agreement by feature-checking and generally accepted analyses (e.g. Carstens, 2000; Ritter, 1991) that posit a functional category Number

Phrase (NumP) between DP and NP, the proposed mechanism for number agreement is the following: given that number features in Romance, and therefore in Spanish, are strong, nouns raise from N to Num, where their uninterpretable [ $\pm$  plural] features are checked against the interpretable number features in Num. This movement yields the N Adj word order, the Adjective Phrase (AP) being located somewhere below NumP. Number features on APs are also checked in Num.

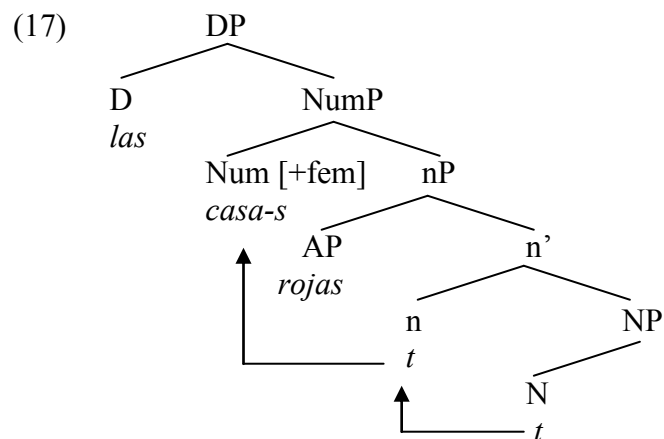


### 3.2. Gender Agreement

Three main accounts describe how gender agreement takes place in languages such as Spanish. According to the first view, gender is lexically determined on nouns (Carroll, 1989; Corbett, 1991; Carstens, 2000). Carstens (2000) argues that the gender feature rises with the noun to NumP, and checks uninterpretable gender features on any adjectives in spec-nP. From here it raises covertly to DP, where it checks the gender feature on D.

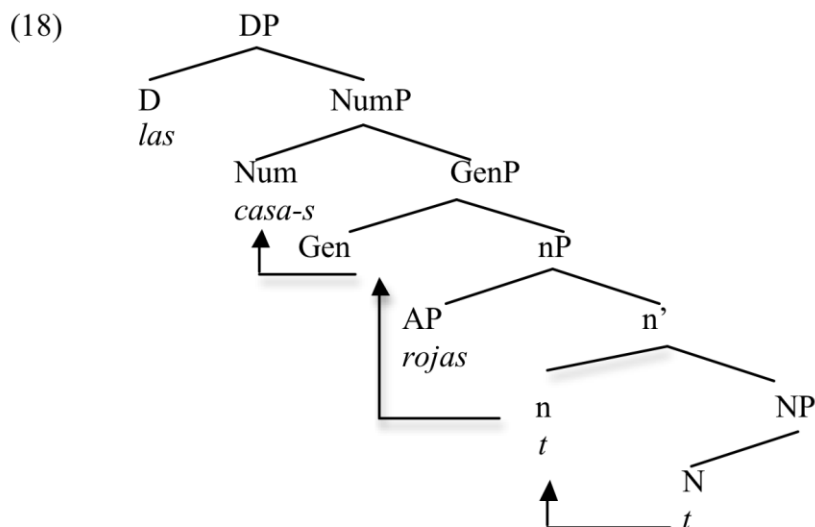


Ritter (1993) proposes an alternative account, arguing that the location of interpretable gender features is subject to parametric variation. For Romance languages like Spanish she posits NumP as the locus of gender features. Thus, in Spanish, the noun raises to Num in order to check both number and gender features.



In a third account, Picallo (1991) proposes that gender projects onto a Gender Phrase (GenP), where nouns check gender features before raising to NumP to check number features.





Despite the fact that these accounts attribute the interpretable gender feature to different categories, the three of them locate gender features somewhere within the DP, where other elements can check their uninterpretable features, thus resulting in the pertinent agreement relations.

#### 4. Approaches to Morphological Variability

Agreement errors are frequent and attested through late stages of acquisition in L2 Spanish (Franceschina, 2002; Hawkins & Chan, 1997; McCarthy, 2008). Given the persistence of such errors in advanced levels of proficiency, the possibility of acquiring a native representation of the features involved in agreement has become a controversial topic. Furthermore, we find disparity of results in the literature, with some studies showing native-like attainment in this area (White et al., 2004; Hopp, 2010), and others illustrating that gender and number agreement constitutes a problem even for highly proficient learners (Franceschina, 2002). This situation raises two main questions, which were briefly discussed in the introduction:

Is gender and number agreement acquirable in the L2 when the features involved are different from or not present in the L1? And, what is the source of agreement errors in advanced second language grammars?

In order to address these questions several theories have been proposed. We can classify them into two groups, which we have briefly explained in the introduction: the representational accounts and the computational or full access accounts. In the following pages we will discuss some of the theories included in each group, and some of the studies that provide evidence for each account.

#### **4.1. Representational Accounts**

As previously mentioned, these accounts posit a deficit in the L2 grammar as the source of morphological variability in L2 learners. The theories included in this first group share the assumptions that L1 and L2 acquisition are different in nature, and that L2 acquisition is largely constrained by the L1, past a critical period (Clahsen & Muysken, 1986; Schachter, 1990; Bley-Vroman, 1989, 1990; Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007). Thus, these theories emphasize maturational constraints, claiming that parameter resetting is not possible after a certain age. As a consequence, late learners cannot acquire native-like abstract representations, resulting in an incomplete grammar. Under representational accounts, this deficit in the grammar is considered to be the root of variability for gender and number agreement.

One of the most discussed theories within this group is the Failed Functional Features Hypothesis (FFFH) (Hawkins & Chan, 1997). This theory specifically claims that functional features like gender are not acquirable in the L2 after puberty, if they are not instantiated in the L1. In order to compensate for the absence of features in the L2, learners are forced to resort to

other mechanisms made available by UG. A proposed mechanism is the adoption of a default—either masculine or feminine in the case of gender—, together with the memorization of exceptions (Hawkins, 2001). Therefore, English-speaking learners of Spanish are expected to exhibit persistent optionality with gender agreement, since they lack the gender feature in their native language.

The Interpretability Hypothesis (Tsimpili and Dimitrakopoulou, 2007), a more recent version of the FFFH, makes different predictions depending on the interpretability or semantic value of features, that is, depending on whether features are interpretable, thus contributing to the meaning of the word, or uninterpretable, therefore not carrying any semantic information (like gender features on determiners and adjectives, in Spanish). The Interpretability Hypothesis makes the specific claim that the problem lies with *uninterpretable* features at the Logical Form (LF) and Phonetic Form (PF) interfaces. Interpretable features, because they carry some semantic value, have a double representation status on the lexicon: a linguistic one, and a conceptual one. Because of this conceptual representation, interpretable features are visible or interpretable at the LF-interface. Thus, these features are not subjected to maturational constraints. This is not the case for uninterpretable features, which lack semantic import, and therefore, are only visible at the PF-interface. The Interpretability Hypothesis predicts that these features will be problematic for late learners at all stages, if the L1 and L2 differ on the setting of parameters concerning those features. In the case of gender and number in Spanish, it is determiners and adjectives that bear the uninterpretable feature. As outlined in the previous section, within the DP, the noun raises to Num, where its interpretable features can check the uninterpretable features on the determiner and the adjective (Carstens, 2000). If uninterpretable features are not acquired, feature checking, and therefore agreement relations become impossible. The Interpretability

Hypothesis then, similar to the Failed Functional Features Hypothesis, predicts variability in gender agreement for English-speaking learners of Spanish.

Some evidence supporting representational accounts comes from a study by Franceschina (Franceschina, 2005), in which she looked at the acquisition of structural case, and gender and number agreement in L2 Spanish. The aim was to examine subjects' performance when features common to both the L1 and L2 were involved, and when features only present in the L2 were implicated, in order to test the predictions of the FFFH. She tested subjects from a variety of L1 backgrounds and arranged them into two groups according to whether their native languages instantiated the optional, uninterpretable feature gender ([+gen] languages), or not ([-gen] languages). All the subjects were very advanced speakers of L2 Spanish. Franceschina gave them one written production task and one comprehension/interpretation task. In the production task the participants had to fill in blanks in sentences with the target pronouns testing case, gender and number agreement, or with distracters such as prepositions and the complementizer *que*, 'that'. In the interpretation task the subjects had to interpret pronouns and adjectives embedded in sentences they read and heard simultaneously, by choosing their referent among three options. This choice was only possible based on gender and number agreement between the pronoun/adjective and the referents, since the three optional referents logically fitted the sentence. Franceschina's prediction was that subjects with [-gen] L1s would show difficulty with gender agreement, while [+gen] L1 participants would perform like native speakers. The results for the production task indicate that case and number were unproblematic for all groups, while gender proved to be much harder for the [-gen] group, satisfying the predictions of Failed Functional Features Hypothesis. As for the interpretation task, there was no difference between the [+gen] group and the native speaker group. On the contrary, there was a significant

difference between the [-gen] group and the native speakers. The difference between the [+gen] and [-gen] groups was marginally significant ( $p = 0.063$ ).

Franceschina concludes that these results fulfill her predictions, and support the FFFH. However, her results are obtained from grouped data, which gives them certain opacity. Some of the [+gen] languages tested, like Italian, have a big overlap with Spanish in gender, even in non-canonical items, while others, like French, exhibit greater differences with Spanish in this respect. In this case, and in order to obtain a clearer picture of the role of the L1, it seems pertinent to have analyses of performance and comparisons between the languages within each group. It would also be relevant to show some examples of individual data. Recall that according to the FFFH it is impossible for [-gen] subjects to achieve native-like attainment in gender agreement. Therefore, even one subject scoring in the native range would be enough to disprove this theory.

#### **4.2. Computational Accounts**

In contrast to representational theories, computational accounts focus on the similarities between first, and second language acquisition. Although they acknowledge that the fact of having already acquired a language by the time we learn a second is an important difference between the two processes, they claim that some basic, key aspects are common to both. Thus, these theories maintain that UG is available in second language acquisition too, and that there is no critical period that impedes its access. Therefore, parameter resetting and native-like attainment are possible, regardless of age of acquisition, and of L1/L2 differences. The main implication relevant to our topic is that variability in agreement in advanced learners is argued

not to be due to incomplete grammars, but to some other issue, which these accounts claim is computational in nature.

One of the theories included in this group is the Full Transfer Full Access (FT/FA) hypothesis (Schwartz & Sprouse, 1996), which, similarly to Failed Functional Features, claims that the L1 is the starting point of L2 acquisition. However, for FT/FA this is only true of initial stages of acquisition, at which point grammatical features are transferred from the L1 to the L2. As proficiency improves, UG becomes available to learners. Therefore, FT/FA, as opposed to Failed Functional Features Hypothesis, does not propose maturational constraints, and claims that parameter resetting, and consequently, native-like attainment, are possible, regardless of age of acquisition and L1/L2 differences. Under this approach, learners whose native language lacks gender agreement still have the potential to acquire it in later stages in acquisition. The main concern for this theory then, is to account for the variability observed in learners.

An explanation for this issue is provided by Missing Surface Inflection Hypothesis (MSIH) (Prévost & White, 2000). According to MSIH, performance is not a reliable indicator of our abstract knowledge or real competence in the L2. This theory claims that native-like attainment is possible, even for late learners. However, even when they have acquired the pertinent L2 features, they may have difficulties accessing them due to performance issues. Specifically, MSIH argues that variability in gender and number agreement is caused by processing or communication pressures, rather than by a deficit in the grammar. Thus, this theory predicts that learners will perform better on untimed tasks, where reduced pressures ease retrieval of the relevant representations, than in spontaneous production or timed tasks.

A study that directly addresses the issue of representational versus computational accounts is White et al., 2004, in which the authors argued against Failed Functional Features,

and in support of Full Access theories. They collected data from French and English-speaking learners of Spanish in one comprehension and two production tasks that tested gender and number agreement. The first production task was an adaptation of the ‘Guess Who’ game aiming at the elicitation of DPs containing adjectives. In this task the subjects had to guess the character that the experimenter had chosen from a series of cards, by asking questions about the physical appearance of the character, such as ‘*Tiene barba roja?*’ ‘Does he have a red beard?’. In the second production task, subjects had to describe three pictures shown to them by the experimenter. The comprehension task was a picture identification task in which subjects were asked to interpret null nominals included in a story. Interpretation of the nominals was based on gender and number agreement between one of three objects represented in pictures, and the determiner and adjective in the null nominal DP. Thus, subjects would read sentences like ‘*Ponlas ahí cerca de la roja.*’ ‘Put them over there by the<sub>FemSg</sub> red<sub>FemSg</sub> [one]’’, and would have to choose the corresponding picture according to gender and number agreement, in this case a feminine, singular object. In order to make sure that the subjects knew the vocabulary items shown in the pictures, along with their gender, they took a vocabulary test including the same pictures, immediately after the comprehension task.

The predictions according to Failed Functional Features (FFFH) and Full Transfer/Full Access (FT/FA) are the following: since English does not instantiate gender features and French does, according to FFFH only the French learners should be able to acquire gender agreement to native-like levels, and therefore, they should outperform the English speakers in all tasks, at all levels of proficiency. The FT/FA hypothesis predicts that the French learners should perform better than the English-speaking learners on gender at the beginner levels, because the L1 English learners are still transferring the L1 features to the L2 at this stage, whereas the French

already have the gender feature in their inventory. However, at advanced levels of proficiency, in contrast to what FFFH predicts, not only the French, but also the English learners should perform in a native manner, since they should have access to UG by then. Because number is present in the three languages, none of the theories predicts problems with number agreement.

The results showed that advanced learners performed like native speakers in both tasks regardless of the L1, contra Failed Functional Features, and in support of Full Access theories. However, the predicted advantage with gender for the low proficiency French group compared to the low English group wasn't found. Both groups performed better with number, and in fact, there were no significant differences between them. These results falsified the Full Transfer part of the theory, which didn't allow for a full support of the FT/FA hypothesis. Nevertheless, the results also challenge the Failed Functional Features Hypothesis (FFFH), since this theory wouldn't predict that the French low proficiency group performed better with number than with gender, or showed equal performance to the English group.

Another relevant finding was that errors were more frequent in one direction, both in production and comprehension. Specifically, masculine seemed to be extended to feminine contexts, as in '*Tiene una barba rojo*' "He has a red<sub>FemSg</sub> beard<sub>MSg</sub>". White et al. categorize this as use of default morphology. This unidirectionality of errors is not predicted by either FFFH, or FT/FA. As an alternate explanation, White et al. suggest that this phenomenon may be compatible with how natural language in general works, and therefore, may be found in native speakers as well. They point out that this option is discussed within the Distributed Morphology framework (Halle & Marantz, 1993), which views the representation of features as asymmetrical (we will review this theory in detail in following sections). Under this approach, the insertion of unmarked or default forms is not ruled out. Some contexts require the insertion of a marked



form, while others require an unmarked form. Normally, both natives and learners supply the form that corresponds to the context. However, learners sometimes supply default forms in contexts requiring a marked form. As an explanation for why learners draw on defaults in these cases, White et al. point at problems accessing the correct form of words due to performance issues, as suggested by Missing Surface Inflection Hypothesis (MSIH) for example. However, they acknowledge that MSIH was proposed on the basis of production data, while in this study the use of defaults was observed both in production and comprehension.

The studies reviewed above investigated L2 gender and number agreement by focusing on the emergence versus the absence of learner variability. However, none of these studies have explored the nature of this variability in depth. In the next section we discuss variability on learners and we review one study essential to our investigation that focuses on the matter.

## 5. Variability on Agreement in L2 Spanish

As briefly explained in the introduction, in Spanish masculine and singular agreement morphemes are frequently used by learners in feminine and plural contexts respectively (example 5 repeated below), as well as their own. In contrast, the use of feminine and plural agreement markers in masculine and singular contexts is not so often attested (example 6).

- (5) La casa es \*antiguo, pero tiene unas vistas \*bonita.  
 The house<sub>FemSg</sub> is old<sub>MSg</sub> but it has some views<sub>FemPl</sub> beautiful<sub>FemSg</sub>.  
 ‘The house is old, but it has some beautiful views.’

- (6) El libro es \*novedosa y tiene un argumento \*ingeniosos.  
 The book<sub>MSg</sub> is innovative<sub>FemSg</sub> and it has a plot<sub>MSg</sub> witty<sub>MPi</sub>.  
 ‘The book is innovative and it has a witty plot.’

Consequently, several studies have characterized learner variability on gender and number agreement as systematic, and reflective of usage of default morphology (McCarthy, 2008; White

et al., 2004; Montrul et al., 2008). However, not many studies have focused on this systematically as a means to explain the causes of morphological variability. One study that directly addresses this question is McCarthy (2008).

McCarthy investigated whether use of default morphology by learners is present in comprehension, and if so, whether it is qualitatively similar to the one observed in production. She contemplates default morphology from the standpoint of an asymmetrical representation of features, as characterized within the Distributed Morphology (DM) (Halle and Marantz, 1993) framework. In order to better understand McCarthy's study, we will review this theory below.

### **5.1. Distributed Morphology**

According to this theory, the syntax is fully specified with features, whereas lexical items are not: features are not represented as binary, but rather as present or absent. Therefore, some morphemes are featurally *specified* (marked), and others are *underspecified* (unmarked<sup>1</sup>/default form). The notion of 'underspecification' implies that all redundant information can be dispensed with from abstract representations (Bobaljik, 2002). Thus, when we have contrastive feature values like masculine versus feminine, or singular versus plural, it is not required that both of them are represented. The underspecified morpheme can be interpreted from the absence of the feature.

In Spanish, masculine and singular are the underspecified forms for gender and number respectively, according to Battistella's (1990) criteria for establishing markedness<sup>2</sup>. Below we provide examples for the criteria relevant to gender and number:

---

<sup>1</sup> Following Harley (1994), Harley & Ritter (2002), and Cowper (2005), McCarthy assumes that markedness determines the specification of features, and matches *unmarked* with *underspecified*.

-*Syntactic distribution* is broader for unmarked items: in Spanish, borrowed inanimate nouns are usually considered masculine, for example *el<sub>MSG</sub> email*, suggesting that masculine is the unmarked form.

-*Indeterminateness* of meaning of unmarked forms, as opposed to the more specific meaning of the marked form: in Spanish the singular form can be used with a plural meaning, but not viceversa; *la margarita es bonita*, ‘the daisy is beautiful’, can either mean that one daisy is beautiful, or that daisies in general are beautiful. This is also an example of syntactic distribution, since the singular is used in a plural context; the use of the plural in a singular context is not possible. This suggests that singular is the unmarked form for number.

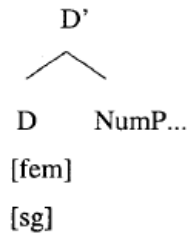
-*Neutralization*, which means that the word loses some specification in certain contexts: in a plural context, the word *padre*, ‘father’, loses the masculine meaning, that is, the word *padres* refers to the father and the mother. This indicates again that masculine is the unmarked form for gender.

The mechanism for agreement proposed in Distributed Morphology is the following. In order to insert the correct morpheme in the syntax, its features need to be checked against those specified in the syntactic context. According to the *Elsewhere Principle* or *Blocking Principle*, the item with the greatest number of matching features should be the one inserted, thus preventing the insertion of an underspecified morpheme in the place of a more specified one. When there is no perfect match of features with any of the available items, the underspecified or *elsewhere* form is inserted (Halle, 1997). This means that the lexical item inserted can be either a *perfect* match, or contain a *subset* of the relevant features, but it cannot represent a *mismatch* or clash of features. Under this approach, the insertion of unmarked or default forms is not ruled

out, since the absence of a feature wouldn't result in a mismatch, but rather in the insertion of a subset of the features specified in the syntax.

(19)

a) *syntactic context*



b) *competing morphemes*

*la* ↔ [fem]

*el* ↔ *elsewhere*

*outcome*

wins competition

blocked

(McCarthy, 2007:19)

To illustrate, in the example above *la* is underspecified for number. However, it contains a subset of the features specified on the syntax, [fem], and it is the morpheme with the greatest number of features matching the syntax. Therefore, by competition, *la* should be the lexical item inserted in this case.

In summary, the *Elsewhere Principle* inhibits the insertion of clashing features, or underspecified items when a more specified item is required and available. If the Elsewhere Principle fails to apply for some reason, two types of agreement errors can occur. When we have a subject DP with a masculine noun, like in example (20), the syntactic context is specified as masculine; none of the available morphemes for gender has a masculine feature, so we should insert the underspecified or *elsewhere* –o. If we insert a morpheme specified as feminine in the agreeing element, in this case an adjective, the result is a clash of features (ex. 21). The same happens for number in (22) and (23), where both subject DPs are specified as singular, but we

insert the plural morpheme –s in the adjective, creating what McCarthy refers to as number ‘feature clash’ errors.

- (20) El libro es rojo. (grammatical)  
 The<sub>MSg</sub> book<sub>MSg</sub> is red<sub>MSg</sub>.
- (21) \*El libro es vieja. (gender feature clash)  
 The<sub>MSg</sub> book<sub>MSg</sub> is old<sub>FemSg</sub>.
- (22) \*La casaa es viejas. (number feature clash fem.)  
 The<sub>FemSg</sub> house<sub>FemSg</sub> is old<sub>FemPl</sub>.
- (23) \*El libro es viejos. (number feature clash m.)  
 The<sub>MSg</sub> book<sub>MSg</sub> is old<sub>MPl</sub>.

In examples (24), (25), and (26) we have the opposite situation. In (24) the context is specified as feminine. The item with the highest number of matching features is the morpheme specified as feminine, *-a*. If we insert *elsewhere* *-o*, there is no clash of features, because the morpheme doesn’t have a gender feature; it is underspecified. However, we still get an agreement error, because the best match available wasn’t inserted. In (25) and (26) the context is specified as plural, but we insert the elsewhere form *-Ø*, which is underspecified. Again, there is no clash of features, but we should have inserted the best match *-s*, which has a plural feature. McCarthy has called these errors ‘underspecification’ errors, also known as default errors.

- (24) \*La casaa es viejo. (gender default)  
 The<sub>FemSg</sub> house<sub>FemSg</sub> is old<sub>MSg</sub>.
- (25) \*Las casaas son vieja. (number default fem.)  
 The<sub>FemPl</sub> houses<sub>FemPl</sub> are old<sub>FemSg</sub>.
- (26) \*Los librosos son nuevo. (number default m.)  
 The<sub>MPl</sub> books<sub>MPl</sub> are new<sub>MSg</sub>.

## 5.2. McCarthy (2008)

McCarthy predicts that learners' errors will be instances of underspecification or default errors, but not feature clash errors. She tested gender and number agreement on intermediate and advanced English-speaking learners and native speakers of Spanish in one production and one comprehension task. The production task aimed at eliciting adjectives and direct object clitics, also inflected for gender and number in Spanish. Participants were shown pictures and were asked questions about them by a native speaker of Spanish. Here is an example of the interactions:

(27)

NS: ¿Qué tiene el chico en la mano?  
What has the boy in the hand?

L2: Una manzana.  
An<sub>FemSg</sub> apple<sub>FemSg</sub>.

NS: ¿Qué va a hacer con la manzana?  
What going to do with the<sub>FemSg</sub> apple<sub>FemSg</sub>?  
'What is he going to do with the apple?'

L2: Va a comerla.  
Going to eat-CL<sub>FemSg</sub>.  
'He is going to eat it.'

In order to elicit adjectives, other questions were asked about the characteristics of the objects in the pictures.

The comprehension task was an adaptation of White et al's (2004) task, where subjects had to interpret clitics embedded in a story by choosing the referent that corresponded to the clitic, among three pictures. The choice could only be done based on gender and number agreement. In order to test knowledge of the vocabulary items in the comprehension task and their gender, subjects took a vocabulary test prior to the task. In this test the participants were

presented with the pictures in the comprehension task, and were asked to write the noun corresponding to each picture and circle the appropriate article according to gender.

The results for intermediate learners showed that variability in gender agreement was present in production, with 82.3% accuracy on clitics and 79.7% on adjectives, as well as in comprehension, with 79.7% accuracy on the interpretation of clitics. Furthermore, the error patterns found in both tasks were qualitatively similar: default errors were more frequent than feature clash errors, that is, masculine was extended to feminine contexts. This resulted in higher accuracy with masculine contexts than with feminine contexts. To illustrate, the means for the intermediate group in the comprehension task were 90.9% in masculine contexts, and 68.4% in feminine contexts. The few errors found on number agreement also pointed to default morphology, with singular extended to plural contexts. This group was significantly different from the native speaker group in both tasks.

As for the advanced learners, there was some evidence of variability in the production of clitics, reflected on higher accuracy on masculine items (100%), versus feminine items (84%). This variability wasn't found on the production of adjectives, where there was no difference between this group and native speakers. With regards to comprehension, some variability was found on the same direction as described above, although performance was better than in production (92% vs. 97.3%), and the learners were not significantly different from the natives.

Largely based on the 79.7% accuracy of the intermediate group in the comprehension task, and the similarity of errors to those found in the production task, McCarthy concludes that contra Missing Surface Inflection Hypothesis, the source of variability must be representational. Her argument is that communication pressures should be at least reduced in comprehension, making variability much less evident on comprehension tasks, if not inexistent, contrary to what

White et al. (2004) and herself found. On the other hand, she doesn't support Failed Functional Features Hypothesis either, as she claims that the representation issue is in the morphology and the representation of features, and not in the syntax.

Instead, McCarthy proposes a developmental pattern in the acquisition of gender features, which ascribes the representational deficit to lower levels of proficiency. According to this pattern, learners first acquire the masculine feature, and lack the feminine, causing errors to be bidirectional, that is, masculine is overextended to feminine contexts, and feminine is overextended to masculine contexts. This was the case for five of her intermediate learners. However, as proficiency increases, the feminine feature is acquired, the asymmetry surfaces, and defaults become systematic; we only see overextension of masculine, as in the rest of McCarthy's intermediate participants, and in the advanced group. This seems to be supported by her results for the production of clitics, where advanced learners used masculine as a systematic default, whereas intermediates were variable on both features. In contrast, in the production of adjectives, both groups showed use of masculine as a default, which McCarthy suggests may be due to the use of masculine adjectives as citation forms. In any case, McCarthy assumes the feature asymmetries proposed by Distributed Morphology for native grammars. Consequently, the default patterns observed in the advanced group reflect a native representation of features, while the bidirectional error patterns observed in some of the learners in the intermediate group suggest a representational deficit.

To sum up, McCarthy argues for representational issues in the L2 morphology at initial stages of development based on two facts: one, the variability she found in comprehension in the intermediate group (80% accuracy), where communication pressures should be reduced, and



therefore can't explain variability; and two, the nature of the errors in comprehension, similar to that in production, with prevalence of underspecification errors over feature clash errors.

The second point, however, contradicts her developmental approach. Since McCarthy claims that default morphology emerges *when the features are acquired*, thus reflecting the asymmetrical representation proposed for native speakers, this pattern can't be taken as evidence for a representational deficit in the intermediate group. Rather, it would entail that natives and learners share the same representation of features, as McCarthy assumes for the advanced group. Therefore, an explanation is needed for why, this being the case, her advanced learners and some of her intermediate learners resort to defaults at all, while native speakers don't.

Another issue with McCarthy's approach is the assumption that variability in comprehension means that there is a deficit in the grammar. Although Missing Surface Inflection Hypothesis initially argued that variability is due to performance factors, which may be reduced in comprehension, many researchers (McDonald, 2006; Hopp, 2010) have argued that performance or communication pressures may be due to *processing issues*, which may be relevant to both production and comprehension. In McCarthy's comprehension task learners had to decompose the morphology of clitics in order to choose a syntactic context, which may require more processing resources than a regular grammaticality judgment task for example, and thus be more taxing for learners.

Given that production and comprehension tasks could entail similar computational effort, manipulating the processing burden of the tasks seems a more suitable methodology to tackle computational difficulty versus representational issues. In the following section we review two studies that implement this methodology with learners and native speakers in order to look for similar outcomes of increased processing burden.

## **6. Variability in Native Speakers**

As we have seen from the literature reviewed so far, the causes of variability in L2 morphology have been the subject of extensive research. We have discussed claims that posit a representational deficit originating in late acquisition as its source. However, evidence from several studies indicates that this might not be the case.

Research from the field of first language acquisition suggests that morphological variability may not be the result of acquiring a second language after a critical period. Hopp (2010) has pointed out that several studies looking at L1 acquisition of German (e.g. Clahsen, Eisenbeiss & Penke, 1996; Eisenbeiss, 1994; Eisenbeiss, 2002; Eisenbeiss, Bartke & Clahsen, 2005/2006; Schütze, 1997) have found variability in children which is qualitatively very similar to that evinced by the L2 learners in one of his studies (Hopp, 2010). The two populations seem to overgeneralize nominative case to accusative contexts, and accusative case to dative contexts. These results suggest that morphological variability is a developmental phenomenon common to both early and late acquisition, and therefore, may not be related to the existence of a critical period for morphology in the second language.

We have also considered an alternative explanation that proposes that learner variability may be the result of computational difficulties in the second language. This hypothesis can be tested by manipulating the processing burden involved in different tasks and exploring its effects on learners' performance. Some studies (Hopp, 2010; McDonald, 2006) have compared learners' performance in online vs. offline tasks, taking advantage of the potential differences of the tasks in terms of processing demands. Online tasks test learners' ability to perform in real time, placing a considerable computational strain on them; on the other hand, in offline tasks there is a gap between the presentation of the stimulus and the learners' response, which diminishes the

processing load involved in the task. The logic behind the use of this methodology is that if learners who show non-native-like patterns under taxing circumstances (online task) also show native-like behavior in a task where computational demands are reduced (offline task), it could be argued that they actually have a native representation of the structure in question, while they may not be able to successfully retrieve it when under increased processing stress (Prévost & White, 2000).

An implication that follows from this proposal is that native speakers might exhibit similar errors to those of learners when exposed to increased computational burden. Some researchers (McDonald, 2006; Hopp, 2010) have approached this question by using different ‘stressors’ with native speakers in order to elicit errors and compare native errors to those made by L2 learners. The purpose of these stressors is to try to recreate the conditions of processing an L2, which are argued to be more extreme than those of processing a native language (McDonald, 2006). With native speakers, the processing burden is increased through the incorporation of an extra load to the task used, such as speed, noise, visual distracters, or a simultaneous secondary task, among others. The idea is that if L2-like error patterns emerge in native speakers, the errors are likely due to the additional processing load integrated in the task, and not to a defective grammar. Again, a similar case could be made for L2 learners: if they make agreement errors on a task involving higher processing difficulty, but they show evidence of native-like behavior on some offline measurement where processing difficulty is reduced, it could be argued that the agreement errors they make may result from the extra computational load involved in processing a second language in real time.

McDonald (2006) tested native speakers and L2 learners of English from 15 different language backgrounds. She conducted two experiments investigating various grammatical

constructions (word order, regular past tense, S-V agreement, plural agreement, etc). In Experiment 1 she administered several independent tasks measuring the subjects' individual processing abilities, including working memory span, decoding ability, and speed of processing. Both native speakers and L2 learners took the tasks in English, as the aim was to measure processing capabilities in the native and second language respectively. In addition, the participants took a grammaticality judgment task testing their performance on the constructions mentioned above. The results showed that the individual processing abilities measured were significantly poorer for the L2ers (in the L2) than for native speakers, suggesting that the efficiency of the processor is reduced when we use a second language. Moreover, a positive correlation was found between the subjects' performance in the grammaticality judgment task, and L2 working memory and decoding ability. Given the fact that correlations do not entail causality, a second experiment was conducted to further investigate the effects of processing difficulties on grammaticality judgments.

In this experiment additional native speaker groups were given different grammaticality judgment tasks containing the same sentences as in Experiment 1. Each task involved the judgment of the sentences under a different 'stress' condition: memory load (low and high), noise, response deadline, and compressed speech. Each condition targeted one of the specific processing capacities previously tested (working memory span, decoding ability, and speed of processing). Results showed that working memory and decoding ability scores correlated with the subjects' performance in the conditions burdening these processing abilities, namely the memory load and noise conditions respectively. Crucially, the results revealed a very similar scale of vulnerability in the constructions tested for stressed native speakers and L2 learners in Experiment 1: articles and regular morphology were the most vulnerable structures, and word

order was the least affected by the stressors. The noise and high memory load conditions yielded the strongest similarity between natives and learners.

Two important conclusions can be drawn from this study: first, that processing abilities are reduced in a second language; and second, that morphology is considerably susceptible to processing stress, both in the first and second languages. These results support the proposal that differences in morphology between native speakers and L2 learners may result from the increased computational difficulty of processing sentences in a second language, which resembles listening to sentences through noise or with a high memory load in the native language.

A more recent study by Hopp (2010) explores the role of computational difficulties, the properties of the L1, and proficiency in the L2, in the incidence of morphological variability in L2 German. In order to look at these variables Hopp tested advanced and near-native speakers from different L1 backgrounds (English, Dutch, and Russian), and native speakers of German. It is important to keep in mind that among the L2ers native languages, Russian is the most similar to German with respect to the morphological properties tested, followed by Dutch and English as the least similar. Four experiments were conducted. Experiments 1 and 2 compared learner variability in online vs. offline tasks, in order to explore the effects of increased processing demands in L2 performance.

In Experiment 1 learners were tested on their offline knowledge of case inflection and word order in German in a grammaticality judgment task. The results for the advanced groups showed that only the L1 Russian subjects behaved like native speakers, suggesting that L1 transfer may still be a factor at the advanced levels of proficiency (recall that Russian was the

most similar to German in terms of morphosyntax). As for the near-native speakers, all of the groups showed native-like responses, regardless of L1-L2 differences.

In Experiment 2 learners were tested on case and subject-verb agreement on a self-paced reading task, allowing for the intended offline vs. online comparison. In contrast to Experiment 1, in Experiment 2 none of the advanced groups showed native-like behavior, indicating that although the advanced L1 Russian have the pertinent grammatical knowledge, as we saw in the untimed grammaticality judgment task, they cannot use it during online sentence processing. Most importantly, in Experiment 2 all near-native groups performed like native speakers in reading slowdowns for the target regions. Nevertheless, reading times were longer for all of the L2 groups than for the native speakers. This suggests that near-native speakers may only be able to use inflection in a native way when they have more time to process the pertinent information. Therefore, in order to test whether speed constrains the target use of inflection, two more experiments were conducted.

Experiments 3 and 4 established a comparison between native speakers under extraordinary computational difficulties and L2 learners. Experiment 3 was taken only by the L2 groups, and tested case and subject-verb agreement using a speeded grammaticality judgment task. This task is a timed version of the regular grammaticality judgment task in which sentences are presented word by word at a specific rate, and subjects are asked to judge sentences in a very limited time window. The base presentation rate for the sentences in this experiment was 250 ms per word (Speed 1) plus 17 ms per character. The results showed that in this higher speed condition only the near-native L1 Russian group could still perform like the native speakers.

Experiment 4 tested native speaker performance on inflection under increased processing pressure, in order to look for potential similarities with the near-natives in Experiment 3. In this

experiment gender marking and auxiliary selection were added to the constructions previously tested. The task was the same as in Experiment 3, with the addition of four more speeds. These were obtained by incrementally reducing the original base presentation rate to 155 ms (Speed 2), 105 ms (Speed 3), 88 ms (Speed 4), and 71 ms (Speed 5), plus 17 ms per letter, and each of the speeds was given to a different group of native speakers. Results showed evidence of considerable breakdown at Speed 5. Critically, there was a striking parallel between this native speaker group's accuracy across sentence types and that of the L1 English and Dutch groups at Speed 1. For both populations case and gender were clearly affected by speed, while grammatical word order stayed pretty accurate. Furthermore, the mentioned L2 groups were statistically indistinguishable from the native group in all sentence types.

From these results Hopp concludes that 'non-target L2 performance at the L2 endstate is partially caused by computational limitations in using inflection, or by lower degrees of efficiency in processing inflection in the L2 compared to the L1' (Hopp, 2010: 924). In addition, based on the results of the advanced and near-native Russian learners in experiments 1 and 3 (respectively), he points at L1 transfer as a factor facilitating or inhibiting computation in the L2, arguing that 'if analogous types of inflection are not used on processing the L1, difficulties in L2 processing appear to be magnified' (Hopp, 2010: 925). Finally, and importantly, based on the English and Dutch performance in the untimed and self-paced reading tasks, Hopp proposes that native-like performance is possible at high levels of proficiency regardless of L1 properties when the computational load is not increased.

The type of designs outlined in the studies above highlight two interesting comparisons: first, L2 learner performance when they are under increased computational pressure compared to when they don't have an additional burden; and second, native speaker performance under

increased computational pressure compared to L2 learner performance. The result is an innovative method to explore the effects of increased processing demands both on first and second languages, and to examine the similarities and differences existing between these two.

## **7. The Present Study**

Building on both McCarthy (2008) and Hopp (2010), the present study tests learner and native sensitivity to default and feature clash violations under different task demands. The uniqueness of the study lies in the incorporation of the methodology used by Hopp (2010) for the first time to explore the nature of variability in gender and number agreement in low, intermediate, and advanced learners of Spanish. Our learners are native speakers of English, a language in which number but not gender is instantiated, in contrast to Spanish, which instantiates both. This approach will allow us to explore feature asymmetries in L2 and native grammars and assess their differences and/or similarities, and thus evaluate representational and computational accounts of morphological variability.

### **7.1 Research Questions and Hypotheses**

Our research questions are formulated in terms of the two main aspects manipulated in the study: the character of the violations tested and the computational demands involved in the tasks. Since both learners and natives are relevant to our study, we will discuss our research questions and hypotheses for each group separately.

#### **7.1.1 L2 Learners**

RQ1: a) Do L1-English learners of Spanish perform better with number than gender, given the absence of *gender* in English?



If, in line with Failed Functional Features Hypothesis (Hawkins & Chan, 1997) and Interpretability Hypothesis (Tsimplici and Dimitrakopoulou, 2007), variability in agreement results from the lack of gender features in the learners' native grammar (and therefore in their L2 grammar as well), we should see an advantage in performance with number over gender *in all proficiency groups, regardless of task demands*.

In contrast, if Full Transfer/Full Access (FT/FA) and Missing Surface Inflection Hypothesis (MSIH) are on the right track, different outcomes are predicted depending on proficiency. Lower levels of proficiency should show the advantage for number, because at this stage they are still transferring from the L1. On the other hand, at high levels of proficiency learners have access to UG, however, they may or may not perform like native speakers, according to MSIH, due to performance issues. Nevertheless, they have the potential to show native-like behavior with both gender and number.

b) Are learners more accurate detecting **feature clash** than **default errors**?

Failed Functional Features Hypothesis (FFFH) and Interpretability Hypothesis predict general variability, not a specific asymmetry in errors across learners. As mentioned in our discussion of these theories, learners might resort to default forms as a strategy, not as a result of feature asymmetries. Therefore, there is optionality in what forms are chosen as defaults. According to FFFH and Interpretability Hypothesis, there should be no consistency in the choice of defaults across learners.

The computational account perspective assumes that systematic variability does not come from a defective L2 grammar, but from difficulties accessing the correct form during online processing. According to Missing Surface Inflection Hypothesis (MSIH), the specific asymmetry tested (higher accuracy with feature clash than default errors) may or may not emerge for

learners who have correctly acquired the relevant features, depending on processing pressures and how well learners can manage them. With regards to lower levels of proficiency, McCarthy (2008) specifically predicts random errors rather than a specific asymmetry for gender, since beginning learners haven't acquired this feature yet.

RQ2: How do different task demands impact learner accuracy and sensitivity to default and feature clash errors?

FFFH and Interpretability Hypothesis predict problems with gender for all learner groups *regardless of task demands*.

Supporting on Missing Surface Inflection Hypothesis, Full Transfer/Full Access predicts potential native-like performance for highly proficient learners in tasks involving low processing demands. In more taxing tasks, use of default morphology is expected to emerge (except for low proficiency learners), and possibly differences with gender and number.

We tested these questions by comparing learner performance on the detection of gender versus number violations, and default versus feature clash violations in an **Untimed** versus a **Speeded Grammaticality Judgment Task (Speeded GJT)**.

### 7.1.2 Natives

RQ3: Is native speaker accuracy impacted by task demands? Are native speakers more accurate detecting **feature clash** than **default errors**?

If morphological variability in L2 learners is due to increased processing difficulty, then native speakers may show declines in accuracy qualitatively similar to those of learners in a task with increased computational demands (Hopp, 2010). That is, native speakers, like L2 learners,

may resort to the underspecified feature under taxing circumstances, and thus show the same feature asymmetries as L2 learners.

If, on the other hand, native speakers and L2 learners have different abstract representations of features, native speakers are not expected to show qualitative similarity (in terms of feature asymmetries) to learners under any circumstances.

We addressed these questions by comparing native performance on the detection of default and feature clash violations in a **Speeded GJT** using three increasing presentation rates. In addition, we compared the quality of errors in native speakers under pressure and L2 learners, in order to identify any potential similarities or differences in the patterns observed.

The answers to our research questions will contribute to the topic of morphological variability and its sources, and in turn, to the broader question of whether features absent from the L1 are acquirable or not.

## **7.2 Method**

### **7.2.1 Stimuli**

We created a total of 216 items, of which 108 were experimental items manipulating agreement between a noun and a predicative adjective; half of them contained a masculine noun, and the other half, a feminine noun. Forty-nine of the total 108 nouns used come from Alemán-Bañón (2010), an ERP study on the processing of gender and number agreement by L2 learners of Spanish. All the adjectives included in the experimental items except two, come from this study as well. Each of the experimental sentences had six versions corresponding to six different conditions. A complete list of our stimuli can be found in Appendix 1. The example below shows the conditions with a masculine noun as the trigger of agreement.

- (28) Juan dijo que vio un colegio que era **antiguo** en Londres. ✓  
 Juan said that he saw a school<sub>MSg</sub> that was old<sub>MSg</sub> in London.
- (29) \*Juan dijo que vio un colegio que era **antiguoA** en Londres. (**gender feature clash**)  
 Juan said that he saw a school<sub>MSg</sub> that was old<sub>FemSg</sub> in London.
- (30) \*Juan dijo que vio un colegio que era **antiguoS** en Londres. (**number feature clash**)  
 Juan said that he saw a school<sub>MSg</sub> that was old<sub>MPi</sub> in London.
- (31) Juan dijo que vio unos colegios que eran **antiguos** en Londres. ✓  
 Juan said that he saw some schools<sub>MPi</sub> that were old<sub>MPi</sub> in London.
- (32) \*Juan dijo que vio unos colegios que eran **antiguo\_** en Londres. (**number default**)  
 Juan said that he saw some schools<sub>MPi</sub> that were old<sub>MSg</sub> in London.
- (33) \*Juan dijo que vio unos colegios que eran **antiguoAs** en Londres. (**number feature clash**)  
 Juan said that he saw some schools<sub>MPi</sub> that were old<sub>FemPi</sub> in London.

The condition exemplified by (28) is **grammatical**, with a masculine, singular adjective agreeing with a masculine, singular noun. In (29) the noun is used in the masculine, singular form again, but the agreeing adjective, although showing the correct singular form, is incorrectly marked as feminine for gender agreement. This condition represents a **gender feature clash** error. In (30) gender agreement is correct, with both elements showing the masculine morpheme. However, the adjective is marked as plural, while the noun shows the unmarked, singular form, creating a **number feature clash** error. Example (31) is **grammatical**, with a masculine adjective in the plural form correctly agreeing with a masculine, plural noun. In (32) both the noun and the adjective display the gender morpheme for masculine; with regards to number, the noun is in the plural, while the adjective is in the singular, constituting a **number underspecification or default** error. In example (33) number agreement is correct, with both elements in the plural form, whereas gender agreement is incorrect, with an adjective marked as feminine mismatching a masculine noun. This condition represents a **gender feature clash** error.

The example below shows the six conditions with a feminine noun as the trigger of agreement:

- (34) Juan dijo que vio una película que era **romántica** en París. ✓  
Juan said that he saw a movie<sub>FemSg</sub> that was romantic<sub>FemSg</sub> in París.
- (35) \*Juan dijo que vio una película que era **romántico** en París. (**gender default**)  
Juan said that he saw a movie<sub>FemSg</sub> that was romantic<sub>MSg</sub> in París.
- (36) \*Juan dijo que vio una película que era **románticaS** en París. (**number feature clash**)  
Juan said that he saw a movie<sub>FemSg</sub> that was romantic<sub>FemPl</sub> in París.
- (37) Juan dijo que vio unas películas que eran **románticas** en París. ✓  
Juan said that he saw some movies<sub>FemPl</sub> that were romantic<sub>FemPl</sub> in París.
- (38) \*Juan dijo que vio unas películas que eran **romántica** en París. (**number default**)  
Juan said that he saw some movies<sub>FemPl</sub> that were romantic<sub>FemSg</sub> in París.
- (39) \*Juan dijo que vio unas películas que eran **románticos** en París. (**gender default**)  
Juan said that he saw some movies<sub>FemPl</sub> that were romantic<sub>MPl</sub> in París.

The grammatical and number agreement conditions (34, 36, 37, and 38) are the same as in the masculine set above. Examples (34) and (37) are the **grammatical** conditions, in the singular and the plural respectively. Example (36) shows the **number feature clash** condition, and example (38) represents the **number default** condition. However, since in this set of examples the noun triggering agreement is feminine, the ungrammatical gender agreement conditions (35 and 39) exemplify **gender default** violations, instead of gender feature clash violations. That is, in this case the adjectives show masculine agreement with a feminine noun. Since half of the experimental items are masculine and the other half are feminine, there is an equal number of items containing feature clash versus default errors.

All of the nouns and adjectives in the present study were controlled for canonicity, in order to avoid testing knowledge of gender, rather than agreement. We also ensured that the

nouns and adjectives used were not subjected to dialectal variation. In addition, we certified that there were no differences in frequency between masculine versus feminine nouns ( $t_{(53)}=.436$ ,  $p=.665$ ), and masculine versus feminine adjectives ( $t_{(53)}=1.521$ ,  $p=.134$ ).

We generated 54 sets of 6 sentences containing masculine nouns, and 54 sets of 6 sentences containing feminine nouns, giving us 9 items per condition. As will be explained in the analysis and results section, we subsequently collapsed across number (sg/pl) conditions for the gender analysis, and across gender (fem/m) conditions for the number analysis, yielding 18 items per condition. Items were distributed across six lists so that each participant saw only one sentence from a given set. The lists were balanced according to number of grammatical and ungrammatical items, and they were randomly assigned to participants. The sentences in each list were automatically randomized at the beginning of the session, and presented using Paradigm Stimulus Presentation by Perception Research Systems Inc. (Tagliaferri, 2005).

A total of 108 distracter and filler items were also created. Seventy-two of these were distracters testing various constructions: 36 of them manipulated Subject-Verb agreement (examples 40-43); 18 assessed the use of the clitic “se” (44-45); 18 tested word order (46-47); 36 were grammatical filler items following two types of structures (48-49).

(40) Hemos visto a unos niños que frecuentemente **juegan** en el parque. ✓  
We have seen some kids who frequently play in the park.

(41) \*Hemos visto a unos niños que frecuentemente **juega** en el parque.  
We have seen some kids who frequently plays in the park.

(42) Hemos visto a un niño que frecuentemente **juega** en el parque. ✓  
We have seen a kid who frequently plays in the park.

(43) \*Hemos visto a un niño que frecuentemente **juegan** en el parque.  
We have seen a kid who frequently play in the park.

(44) Mi vecino Pedro patina con sus amigos en el parque. ✓  
My neighbor Pedro skates with his friends in the park.

- (45) \*Mi vecino Pedro **se** patina con sus amigos en el parque.  
My neighbor Pedro **se** skates with his friends in the park.
- (46) El martes María compró una casa de piedra para su hijo. ✓  
On Tuesday María bought a house made of stone for her son.
- (47) \*El martes María compró una casa de piedra su hijo para.  
On Tuesday María bought a house made of stone her son for.
- (48) Veo que puedes cantar las canciones que destrozaste en el recital.  
I see that you can sing the songs that you destroyed in the recital.
- (49) Me encanta el suéter que tejiste para el invierno de Canadá.  
I love the sueter that you knitted for the winters in Canada.

### 7.2.2 Tasks

The sentences above were presented in two different task formats. Following Hopp (2010), we created these formats by manipulating the conditions under which participants were asked to supply grammaticality judgments targeting agreement. Table 1 below shows the distribution of the participants among the tasks. Three separate groups of Spanish natives took a Speeded Grammaticality Judgement Task (Speeded GJT) in which sentences were presented word by word at three different rates of presentation (Speeds 1, 2, and 3). The details of this task will be discussed in the following section. Three proficiency groups of L2 learners (low, intermediate, advanced) took either a Speeded Grammaticality Judgment Task (Speed 1) or an Untimed Grammaticality Judgment Task (Untimed GJT) that included the same sentences. This design allowed us to examine the quantitative and qualitative effects of speed on native speakers' and L2 learners' sensitivity to default versus feature clash errors, as well as any similarities between the groups. Both the tasks and the participant groups are explained in detail below.

### **7.2.2.1 Speeded Grammaticality Judgment Task (GJT)**

Following Hopp (2010) and previous studies (Meng & Bader, 2000; Bader & Meng, 1999; Schlesewsky & Frisch, 2003; Vogel & Frisch, 2003) we used a speeded grammaticality judgment task (speeded GJT) to measure responses under a high processing load. The task tested the subjects' ability to rapidly detect and judge gender and number default and feature clash errors presented at an elevated processing speed. The participants had to read sentences in Spanish on a computer screen and judge their grammaticality by pressing 'bien' or 'mal' (good/bad) on the keyboard. They were instructed to respond as quickly and as accurately as possible. In order to do this more effectively they were also asked to rest their hands on the keyboard throughout the experiment. The experiment started with ten practice items. The participants received feedback on their answers, and they were warned that this would not be the case in the actual experiment. The purpose of this practice was that they familiarized themselves with the procedure and that they understood what we meant by a good and a bad sentence in Spanish. The sentences were presented in the middle of the screen after a fixation star, one word at a time, in 14 pt. Microsoft Sans Serif font over a blue background. After the sentence was presented the background color changed to green and the words 'Bien' 'Mal' appeared on the screen, indicating that this was the time to judge the sentence. As soon as the judgment was given the fixation star appeared on the screen again and the participants could rest or move on to the next sentence by pressing the spacebar. In addition, 5 breaks were offered throughout the experiment.

In order to manipulate processing demands we used 3 different speeds or presentation rates: 155 ms per word for Speed 1; 71ms per word for Speed 2; and 37 ms per word for Speed



<sup>2</sup>. Seventeen milliseconds per character were added to each of the base presentation rates in order to adjust the time of presentation of the words to their length. Since the critical words (the adjectives) in our stimuli were three to twelve characters long, this means that in Speed 1 the critical words were presented on the screen for 206 ms to 376 ms; in Speed 2, the range of presentation was from 122 ms to 292 ms; and in Speed 3, the range was from 88 ms to 258 ms<sup>3</sup>. In order to make the task more taxing we also introduced a visual distracter consisting of 12 hash marks (#####), which was presented for 50 ms in between words. To decide on the length and presentation rate of the distracter we took into account three factors: first, that most of our adjectives were quite long, which made them very salient; second, that at such a high speed small words like prepositions were really hard to perceive; and third, we didn't want the distracter to slow down the presentation rate. We chose a length of 12 hash marks because it disfavored long words like adjectives, as it approximated their own length, and benefited smaller words, as they became more salient when framed between these symbols. In addition, the 50 ms presentation time was enough for the mask to be distracting, and not so much that it reduced the overall pace of the task.

---

<sup>2</sup> We selected these speeds based on previous studies using the task (Hopp, 2010; Meng & Bader, 2000; Bader & Meng, 1999; Schlesewsky & Frisch, 2003; Vogel & Frisch, 2003), and our own pilot study testing learners and native speakers. Hopp's (2010) L2 subjects evinced breakdown at 250 ms per word plus 17 ms per character, while the native speakers' breakdown was at 71 ms per word plus 17 ms per character, for case and subject-verb agreement in German. The results of our pilot study showed that these speeds weren't demanding enough to elicit breakdown in gender and number agreement in Spanish. Thus, we tested different speeds obtained by repeatedly subtracting 17 ms to the base presentation rate, following Hopp (2010), until a breakdown was observed.

<sup>3</sup> Since plural words are one character longer than singular words (they add an -s), they were presented for 17 ms longer than singular words. Nevertheless, our results do not reflect this difference, given that the effects observed for number in the speeded task are also present in the untimed task in the low proficiency group (see results section).

### 7.2.2.2 Untimed Grammaticality Judgment Task (GJT)

In this version of the task the whole sentence was presented in the middle of the screen. Subjects decided when to move from one sentence to the next and when to give their judgment by pressing the spacebar. Therefore, there was no time limit to read or judge the sentences, and no visual distracter. This task was given to L2 learners in order to examine how processing pressure affected them by comparing their performance when they were pressure-free (Untimed GJT) to when they were under stressful circumstances (Speeded GJT).

### 7.2.3 Participants

In this section we describe the L2 and native speaker groups tested in the present study. Table 1 below shows the distribution of the participants among the tasks explained in the previous section.

**Table 1. Participants by task**

<b>Tasks</b>  <b>Groups</b>	<b>Speeded GJT</b>			<b>Untimed GJT</b>
	<b>Speed 1</b> (155ms/word +17ms/character)	<b>Speed 2</b> (71ms/word +17ms/character)	<b>Speed 3</b> (37ms/word +17ms/character)	
<b>Natives</b>	n=12	n=36	n=36	—
<b>L2</b>				
<b>Low</b>	n=16	—	—	n=7
<b>Int.</b>	n=16	—	—	n=4
<b>Adv.</b>	n=15	—	—	n=4

#### 7.2.3.1 L2 Learners

The L2 learners (n = 62) were recruited from Spanish classes ranging from the 200-level to Graduate Teaching Assistants at the University of Kansas. Three learners were recruited from outside the university. Two of them were former students who were still using Spanish for their jobs and/or everyday life, and one of them had had no formal training in Spanish, but had lived

in Honduras for 9 years. All of them were native speakers of English with no substantial exposure to Spanish or other Romance language before the age of 12. On this basis we assume that they started acquiring Spanish past any proposed critical period. Subjects were grouped as low, intermediate, and advanced, according the MLA/DELE (Montrul, 2005) Spanish proficiency test. This is a fifty-item written test that includes a multiple choice and a cloze test assessing vocabulary and grammatical knowledge. This test that has been used in a number of second language acquisition studies on gender and number agreement (McCarthy, 2007; Montrul, 2005; Montrul, 2008; White et al., 2004). Learners with scores ranging from 0-29 are characterized as low proficient; learners scoring between 30-39 are described as intermediate proficient; and learners with scores between 40-50 are categorized as advanced proficient.

Tables 2 and 3 report the L2 groups in the present study along with the mean age, proficiency score, and mean age of acquisition for each proficiency group in each task.

**Table 2. L2 learner characteristics by proficiency group: Speeded GJT**

Task	Proficiency Group	Minimum	Maximum	Mean	Std. Deviation	
Speeded GJT	Low N = 16	Age	19	22	20.38	1.02
		Age of Acquisition	7	20	14.19	2.64
		Proficiency Score	15	29	22.19	4.26
	Intermediate N = 16	Age	18	30	21.75	2.91
		Age of Acquisition	10	17	14.50	2.07
		Proficiency Score	30	38	35.19	2.64
	Advanced N = 15	Age	19	42	26.33	5.80
		Age of Acquisition	9	22	15.00	3.66
		Proficiency Score	40	50	43.87	2.67

**Table 3. L2 learner characteristics by proficiency: Untimed GJT**

<b>Task</b>	<b>Proficiency Group</b>		<b>Minimum</b>	<b>Maximum</b>	<b>Mean</b>	<b>Std. Deviation</b>
<b>Untimed GJT</b>	Low N = 7	Age	18	22	19.43	1.40
		Age of Acquisition	10	16	13.43	2.07
		Proficiency Score	19	23	21.14	1.68
	Intermediate N = 4	Age	18	21	19.50	1.29
		Age of Acquisition	14	14	14.00	.00
		Proficiency Score	30	37	34.00	3.56
	Advanced N = 4	Age	25	28	26.75	1.26
		Age of Acquisition	12	15	13.50	1.29
		Proficiency Score	40	49	45.75	4.03

### 7.2.3.2 Native Speakers

The native speakers in Speeds 1 and 2 ( $n = 72$ ) were recruited at the Universidade de Santiago de Compostela in Galicia, Spain. Thirty-six participants took the experiment in the 37ms condition (Speed 3), while another 36 took it in the 71ms condition (Speed 2). The vast majority were undergraduate students at the university. For the 155ms condition (Speed 1) data were collected from 12 participants who were recruited among the researcher's family and friends. They all had studied at the university level, and had different occupations at the time of testing.

All of our native speakers were bilingual speakers of Spanish and Galician and reported having received education in both languages at different points in their life. Most of the participants reported knowledge of two other languages, generally English as the second, and

French, German, or Italian as the third. None of these characteristics was considered as excluding from the experiment, since all of the subjects were born in Galicia, Spain, where both languages are spoken and used in education, with an advantage for Spanish, mostly in cities. Furthermore, Spanish and Galician are very similar with regards to the gender and number of the nouns we used.

In addition, 3 native speakers were recruited at the University of Kansas to replace three native speakers from Spain that were excluded from the study. Their countries of origin were Cuba, Mexico, and Costa Rica. No differences were observed in their performance with respect to the natives from Spain.

Table 4 reports the native speaker groups along with their mean age at the time of testing for each speed condition.

**Table 4. Native speaker characteristics by speed group**

Speeded GJT	N	Age	
		Mean	Std. Deviation
Speed 1 (155 ms)	N = 12	27.25	2.45
Speed 2 (71 ms)	N = 36	20.67	2.75
Speed 3 (37 ms)	N = 36	20.53	2.43

#### **7.2.4 Procedures**

Subjects first signed a consent form and completed a language background questionnaire. For L2 learners this questionnaire inquired about onset of acquisition, years of study, experience living abroad, use and amount of exposure to the language, other languages spoken etc. They were also asked to rate their skills in Spanish. Native speakers were asked to rate the dominance

of Spanish and Galician in their case, and point out which one was used throughout their education from elementary to university level. Following the completion of the questionnaire, the subjects took the experiment on the computer. The L2 speakers took the written proficiency test following the experiment. The session for the natives ended when they finished the experiment.

The L2 participants were tested either as a group, or individually in computer labs at the University of Kansas. Payment for their participation differed depending on how they were recruited: some of the learners took the experiment as part of their Spanish class curriculum and received no payment; others were recruited from conversational classes and received extra credit for their participation; the rest of the learners were recruited through flyers or department email and were paid \$10 for their participation.

The native speakers recruited at the Universidade de Santiago de Compostela were tested in groups of 12 in the computer lab of the Facultade de Filoloxía, in Galicia, Spain, and they were paid five euro for their time. All other native participants were tested individually in quiet rooms. The three participants recruited at the University of Kansas were tested individually in the Second Language Acquisition lab at the University of Kansas.

## **8. Analysis and results**

In this section we report the analysis and results for all native speaker speed groups and all L2 learner proficiency groups.

We conducted our statistical analysis using  $d'$  scores, which allow to explore differences between acceptance rates for grammatical versus ungrammatical sentences in each subject, unveiling any possible biases towards one response or the other. A  $d'$  score is obtained from the difference in the average acceptance rates between each ungrammatical condition and its

grammatical counterpart<sup>4</sup>, thus collapsing two conditions into a unique score. A d' score near zero represents performance at chance, while perfect performance results in a d' score of approximately 4.0. A negative score indicates that grammatical sentences are treated as ungrammatical or vice versa.

Therefore, we obtained acceptance rates for each subject, from which d' scores were calculated. These were averaged for each condition, for each participant. A mixed repeated-measures ANOVA including the three L2 groups on the Speeded GJT was performed on the resulting data. Another mixed repeated-measures ANOVA was conducted for the three native speakers groups. Follow-up analyses were also conducted for each proficiency and each speed group. The results will be presented separately for L2 learners and native speakers, both descriptively and statistically according to the ANOVAs conducted. Statistical analyses weren't performed on the Untimed GJT data due to small sample sizes. Therefore, these results will be reported only descriptively.

Table 5 below summarizes the conditions we were testing in our analyses.

**Table 5. Noun-Adjective agreement errors tested**

Feature		Default	Feature Clash
Gender	Masculine		colegio <sub>MSg</sub> -*antigua <sub>FemSg</sub>
			colegios <sub>MPI</sub> -*antiguas <sub>FemPI</sub>
	Feminine	tela <sub>FemSg</sub> -*fina <sub>MSg</sub>	
		telas <sub>FemPI</sub> -*finas <sub>MPI</sub>	
Number	Singular		colegio <sub>MSg</sub> -*antiguos <sub>MPI</sub>
			tela <sub>FemSg</sub> -*finas <sub>FemPI</sub>
	Plural	colegios <sub>MPI</sub> -*antiguo <sub>MSg</sub>	
		telas <sub>FemPI</sub> -*fina <sub>FemSg</sub>	

<sup>4</sup> Formula on linked spreadsheet for a Forced-Choice paradigm:  $=(\text{NORMSINV}(H)-\text{NORMSINV}(FA))/\text{SQRT}(2)$ ; H = Hit rate for choosing A when A; FA = False alarm rate for choosing A when B.

Although we didn't predict any differences between gender errors in the singular versus gender errors in the plural, we included both singular and plural noun phrases in the stimuli in order to be comprehensive. However, for our analyses we collapsed conditions across number, thus obtaining one average for gender default (singular and plural), and one average for gender feature clash (singular and plural). Similarly, we didn't predict any differences between number errors in the feminine versus number errors in the masculine, but again, we included both feminine and masculine noun phrases in the stimuli for the purpose of being comprehensive. For the statistical analyses we collapsed these conditions across gender, resulting in one average for number default (masculine and feminine), and one average for number feature clash (masculine and feminine).

## **8.1 L2 Learners**

We will report first overall results including the three proficiency groups, followed by separate analyses for each group.

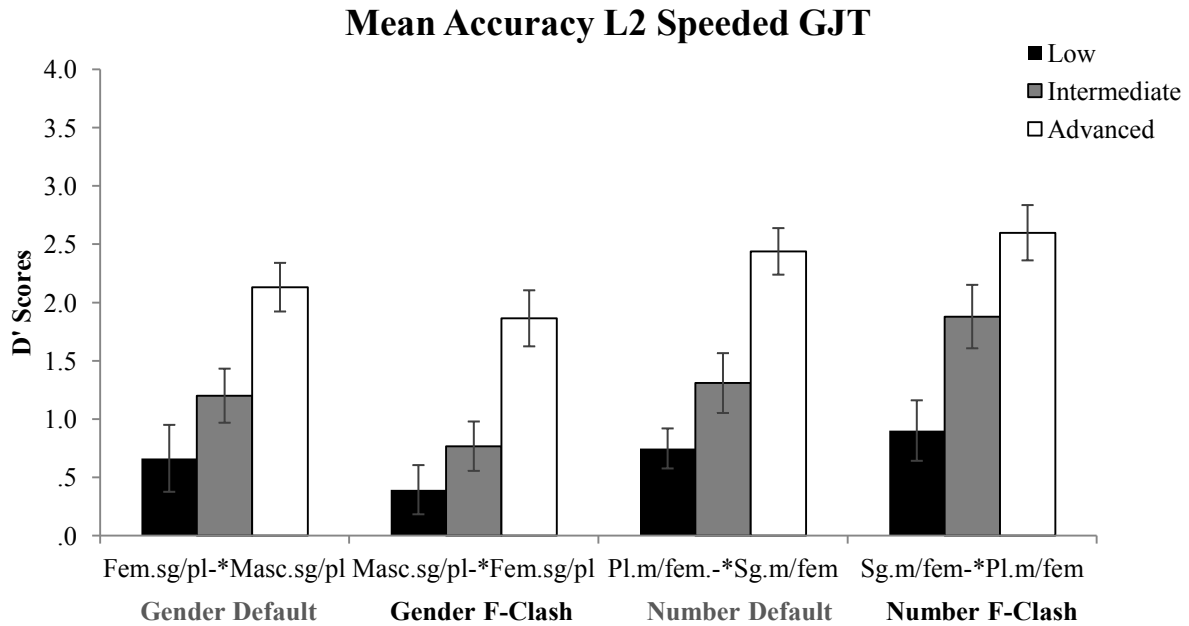
### **8.1.1 Speeded GJT**

Recall that a  $d'$  score near 4.0 indicates perfect performance, while a score of approximately 0.0 represents performance at chance. In this task we do not expect any group to score close to 4.0, because of increased task demands. For the same reason, we expect the low proficiency group to have a mean closer to 0.0.



### 8.1.1.1. Overall Results

Figure 1 represents mean  $d'$  scores for sensitivity to gender and number errors for low, intermediate, and advanced groups in the Speeded GJT.



**Figure 1. Mean  $d'$  scores for L2 sensitivity to gender and number default and feature clash errors across proficiencies in the Speeded GJT.**

We conducted a mixed repeated-measures ANOVA with Feature (*Gender vs. Number*) and Error (*Default vs. Feature Clash*) as within-subjects factors, and Proficiency (*Low, Intermediate, Advanced*) as the between-subjects factor .

Results showed a significant effect for Group,  $F(2,44)=15.48, p<.01$ . Pairwise comparisons showed no differences between the low and intermediate groups, ( $p = .11$ ); the advanced group was significantly better (mean = 2.26) than both the low (mean = 1.29), and intermediate (mean = .68) groups ( $p<.01$ ). A significant effect of Feature was also present,  $F(1,44)=40.36, p<.01$ , reflecting higher sensitivity to Number (mean=1.65) than Gender

(mean=1.17) errors. There were no Error effects in any of the groups. However, there was a significant interaction of Error\*Feature,  $F(1,44)=13.37$ ,  $p<.01$ . This interaction corresponds to higher sensitivity to Default than Feature Clash conditions in *gender*, and better performance with Feature Clash than Default conditions in *number*.

In order to more closely explore the nature of these effects, separate analyses were conducted for each proficiency group.

### 8.1.1.2 Group Results

We conducted separate 2X2 repeated-measures ANOVAs for each of the proficiency groups. The within-subjects factors were Feature (*Gender vs. Number*) and Error (*Default vs. Feature Clash*). The results of these analyses are presented in Table 6 below.

**Table 6. Repeated-measures ANOVAs on learner d' scores for Feature and Error**

Analysis for Feature and Error								
Groups	n	d.f.	Error		Feature		Error*Feature	
			F	(p)	F	(p)	F	(p)
Low	16	1,15	.42	.53	3.61	.08	3.05	.10
Intermediate	16	1,15	.12	.73	40.97	.00	11.81	.00
Advanced	15	1,14	.17	.68	15.91	.00	1.57	.23

Feature (*Gender vs. Number*) analysis: we found a significant main effect for intermediate,  $F(1, 15)= 40.97$ ,  $p<.01$ , and advanced learners,  $F(1, 14)= 15.91$ ,  $p<.01$ , which corresponds to higher sensitivity to *number* than *gender* errors in both groups. The effect for the low proficiency group was marginal,  $F(1, 15)= 3.61$ ,  $p=.08$ , but in the same direction.

Error (*Default vs. Feature Clash*) analysis: there were no significant main effects in any of the groups. However, there was a significant interaction of Error\*Feature for the intermediate

group,  $F(1, 15) = 11.81, p < .01$ , which reflected higher sensitivity to Default (mean=1.2) than Feature Clash (mean=.77) errors *in gender*. For *number*, the reverse pattern, higher sensitivity to Feature Clash (mean=1.88) than Default errors (mean=1.31), was marginally significant ( $t_{(15)} = -1.98, p = .06$ ). This interaction was not significant for the low and advanced groups.

In summary, the three proficiency groups were more sensitive to number than gender errors. In addition, in the intermediate group opposite patterns of errors emerged within each feature: in number, learners were more accurate detecting Feature Clash (50c and 50d) than Default errors (50a and 50b), and surprisingly, in gender, they were more accurate detecting Default (51a and 51b) than Feature Clash errors (51c and 51d).

#### (50) Number

a. Julio dijo que vio unas **telas** que eran **finas**... (number default fem.)  
Julio said that he saw some fabrics<sub>FemPl</sub> that were fine<sub>FemSg</sub> ...

b. ...unos **colegios** que eran **antiguos**... (number default masc.)  
...some schools<sub>MPl</sub> that were old<sub>MSg</sub> ...

Ⓒ ...una **tela** que era **fina**S... (number feature clash fem.)  
... a fabric<sub>FemSg</sub> that was fine<sub>FemPl</sub> ...

Ⓓ ...un **colegio** que era **antiguo**S... (number feature clash masc.)  
... a school<sub>MSg</sub> that was old<sub>MPl</sub> ...

#### (51) Gender

Ⓐ Julio dijo que vio una **tela** que era **fin**O en París. (gender default sg.)  
Julio said that he saw a fabric<sub>FemSg</sub> that was fine<sub>MSg</sub> in Paris.

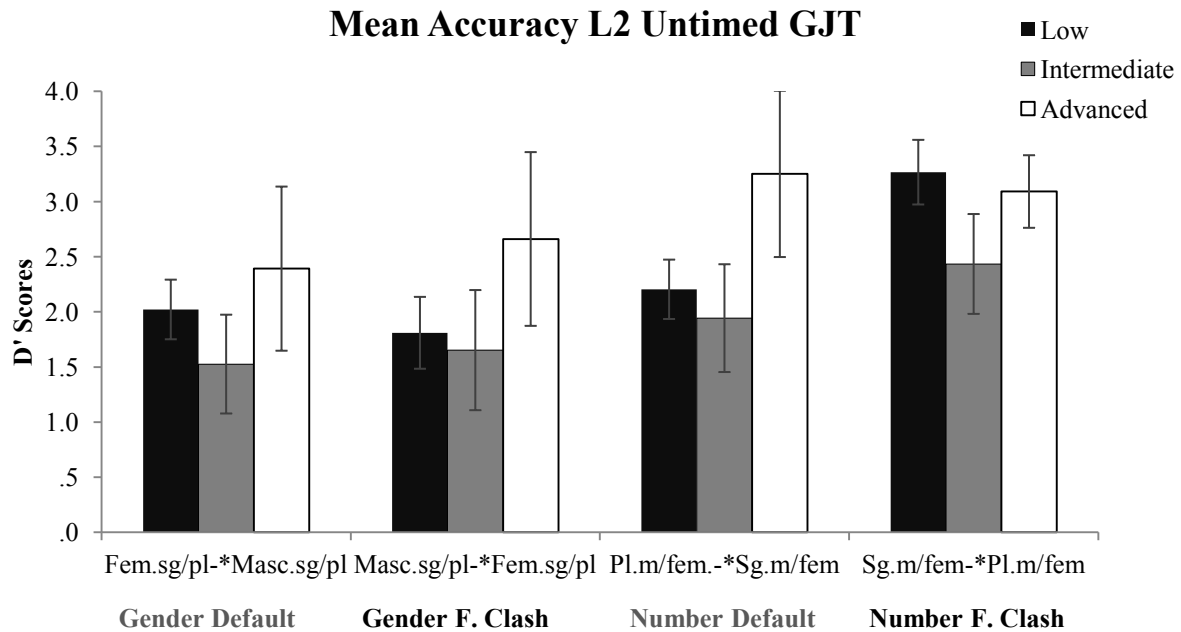
Ⓑ ...unas **telas** que eran **fin**Os... (gender default pl.)  
...some fabrics<sub>FemPl</sub> that were fine<sub>MPl</sub> ...

c. ... un **colegio** que era **antiguo**A ... (gender feature clash sg.)  
... a school<sub>MSg</sub> that was old<sub>FemSg</sub> ...

d. ...unos **colegios** que eran **antiguos**As... (gender feature clash pl.)  
...some schools<sub>MPl</sub> that were old<sub>FemPl</sub> ...

### 8.1.2 Untimed GJT

Figure 2 represents mean  $d'$  scores for sensitivity to gender and number errors for low, intermediate, and advanced groups in the Untimed GJT.



**Figure 2. Mean  $d'$  scores for L2 sensitivity to gender and number default and feature clash errors across proficiencies in the Untimed GJT.**

In this task the scores for the low and intermediate groups were around the middle part of the  $d'$  scale, while advanced learners scored closer to the upper end of the scale. Again, accuracy seemed to be impacted by proficiency. However, due to the small sample sizes of our groups a lot of variability can be observed within each group, with individuals at the low level of proficiency scoring as high as 4.0 in the number feature clash condition, and one individual in the advanced group scoring as low as .38 in the gender feature clash condition. Nevertheless, it is important to notice that two of our advanced learners (out of four) showed perfect performance in every condition.

With regards to differences in performance with gender and number, there seems to be an advantage for number in the low proficiency group, while no differences seem evident in the intermediate and advanced groups. As for differences in sensitivity to Default versus Feature Clash errors, in the low proficiency group the means for Default errors were unexpectedly higher than those for Feature Clash errors in *gender*. In contrast, in *number*, the means for Feature Clash errors were higher than the means for Default errors. In the intermediate and advanced groups there were no apparent differences between Default and Feature Clash conditions in gender or number.

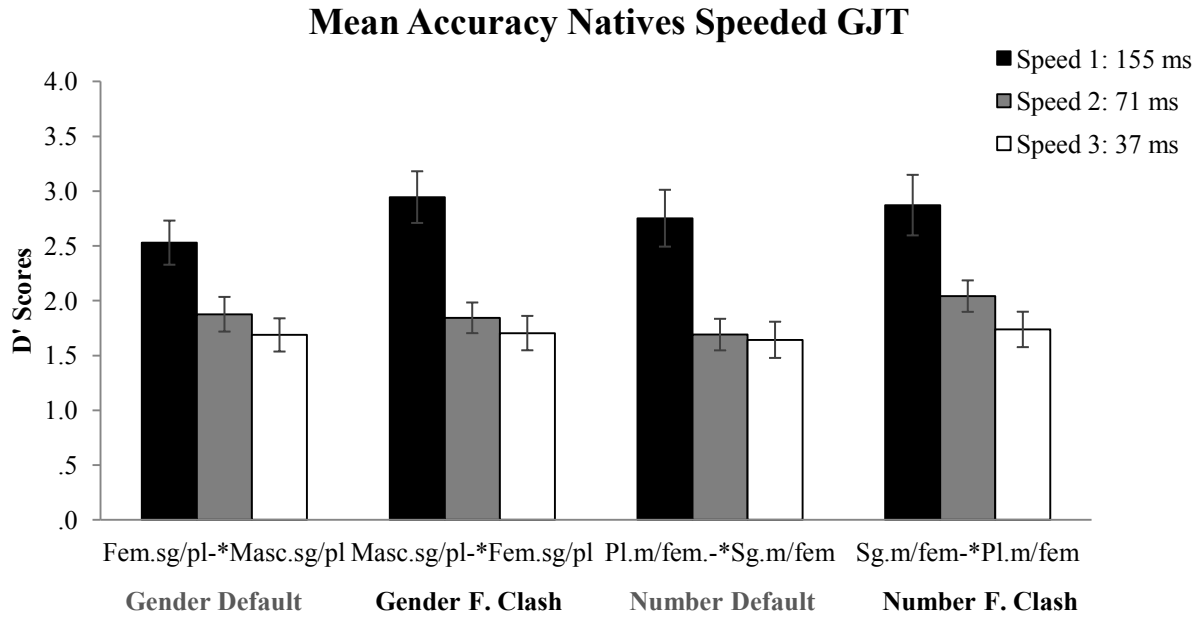
To sum up, in the untimed task the three proficiency groups had higher means than in the speeded task, suggesting an impact of task demands on accuracy. The advantage for number over gender seemed to hold only in the low proficiency group. In this group the same Default versus Feature Clash differences observed for the intermediate group in the speeded task were now apparent. In the intermediate and advanced groups no differences in the detection of Default and Feature Clash errors were evident.

## **8.2 Native Speakers**

We will report first overall results including the three speed groups, followed by separate analyses for each of the groups.

### **8.2.1 Overall Results**

Figure 3 represents mean  $d'$  scores for sensitivity to gender and number errors for the three speed groups.



**Figure 3. Mean  $d'$  scores for native sensitivity to gender and number default and feature clash errors across speeds**

We conducted a mixed repeated-measures ANOVA with Feature (*Gender vs. Number*) and Error (*Default vs. Feature Clash*) as within-subjects factors, and Speed (*Speed =250 ms; Speed 2=71 ms; Speed 3=37 ms*) as the between-subjects factor.

The results showed a significant effect of Speed,  $F(2,81)=8.97, p<.01$ . Pairwise comparisons revealed no differences between the Speed 2 and Speed 3 groups ( $p = 1.00$ ). The Speed 1 group (mean =2.78) was significantly more accurate than both the Speed 2 (mean =1.86) and Speed 3 (mean =1.69) groups ( $p<.01$ ). No Feature effect was present. However, there was a marginal effect of Error,  $F(1,81)=3.71, p<.058$ , reflecting better performance with Default than Feature Clash conditions. We found no significant interactions in these data.

Given these results, separate analyses were conducted for each of the speed groups in order to further investigate the effects found.

### 8.2.2 Group Results

Three separate 2x2 repeated-measures ANOVAs were conducted according to speed group. The within-subjects factors were Feature (*Gender vs. Number*) and Error (*Default vs. Feature Clash*). The results of these analyses are presented in Table 7.

**Table 7. Repeated-measures ANOVAs on native speaker  $d'$  scores for Gender vs. Number**

Analysis for Gender vs. Number								
Groups	n	d.f.	Error		Feature		Error*Feature	
			F	( <i>p</i> )	F	( <i>p</i> )	F	( <i>p</i> )
Speed1 (155 ms)	12	1,11	.95	.35	.38	.55	.75	.41
Speed 2 (71 ms)	36	1,35	2.30	.14	.01	.92	4.36	.04
Speed 3 (37 ms)	36	1,35	.31	.58	.01	.93	.20	.66

Feature (*Gender vs. Number*) analysis: there were no significant main effects or interactions for any of the speed groups.

Error (*Default vs. Feature Clash*) analysis: no significant main effects emerged. However, in the Speed 2 group there was a significant interaction of Error\*Feature,  $F(1, 35)=4.36$ ,  $p=.04$ , which corresponds to higher sensitivity to *number* feature clash (mean= 2.35) than *number* default errors (mean= 1.89), opposed to no difference between *gender* feature clash (mean= 2.45) and *gender* default errors (mean=2.11). This interaction did not emerge in speed groups 1 and 3.

In summary, native speakers showed no differences in the detection of Gender versus Number errors. With regards to sensitivity to the different error types, native speakers in Speed 2 were more sensitive to Feature Clash (50c and 50d) than to Default (50a and 50b) errors.

## (50) Number

- a. Julio dijo que vio unas **telas** que eran **finas**... (number default fem.)  
Julio said that he saw some fabrics<sub>FemPl</sub> that were fine<sub>FemSg</sub> ...
- b. ...unos **colegios** que eran **antiguos**... (number default masc.)  
...some schools<sub>MPl</sub> that were old<sub>MSg</sub> ...
- c. ...una **tela** que era **finas**... (number feature clash fem.)  
... a fabric<sub>FemSg</sub> that was fine<sub>FemPl</sub> ...
- d. ...un **colegio** que era **antiguos**... (number feature clash masc.)  
... a school<sub>MSg</sub> that was old<sub>MPl</sub> ...

## 9. Discussion

The present study explored the nature of variability in gender and number agreement in native and non-native Spanish. Our research was directed to test two possible accounts of this variability: representational accounts and computational accounts. For our purpose we studied three important variables: we examined differences in the detection of Gender agreement violations versus Number agreement violations; we investigated within-feature variability by looking at differential sensitivity to Default errors versus Feature Clash errors in each feature; and we tested how different task demands impact the variable detection of these error types. The following research questions were posed:

### L2 Learners

- RQ1: a) Do L1-English learners of Spanish perform better with **Number** than **Gender**, given the absence of *gender* in English?
- b) Are learners more accurate detecting **Feature Clash** than **Default** errors?



RQ2: How do different **task demands** impact learner accuracy and sensitivity to both error types?

## Natives

RQ3: Is native speaker accuracy impacted by task demands? Are native speakers more accurate detecting **feature clash** than **default errors**?

In this section we will discuss our findings in relation to these questions and to the theoretical accounts tested.

### 9.1 L2 Learners

In our first research question we asked whether learners would show an advantage in performance for Number over Gender, due to the absence of uninterpretable gender in their native feature inventory. Our results showed that this was the case for the three proficiency groups in the speeded task. This difference was marginal for the low proficiency group, which may be explained by their performance at chance in this task.

In the untimed task, however, the advantage for number did not emerge for the intermediate and advanced groups, while becoming more evident in the low proficiency group. This already answers our second research question, suggesting that *when the computational load is not increased*, it becomes possible to overcome L1 transfer at advanced levels of proficiency (Hopp, 2010). In the case of the low proficiency group, reduced task demands resulted in a more robust advantage for number. This seems a logical outcome for this group in this task. Since learners had a better chance to compute agreement, as indicated by their improved performance, existing L1 transfer effects became clearer than when the learners were performing at chance.

These results perfectly match Full Transfer/Full Access. In the absence of an additional processing load, intermediate and advanced learners show equal performance with gender and number features, regardless of the presence of the feature in their L1. At this point, learners have full access to UG and the L2 features. However, the low proficiency group shows signs of transfer, performing more accurately with the feature present in the L1 inventory, namely number.

A crucial result in the learner groups is the fact that intermediate and advanced learners show an advantage with number in the speeded task but not in the untimed task, which is consistent with the predictions by Missing Surface Inflection Hypothesis (computational accounts). Recall that Failed Functional Features Hypothesis and Interpretability Hypothesis (representational accounts) do *not* predict native-like performance with gender *in any task or proficiency group*. In addition, two of our advanced learners did show native-like performance in their least demanding task (Untimed GJT).

Our first question also addressed a potential advantage for detecting Feature Clash over Default errors. This was actually the pattern we found for intermediate learners in the speeded task, and for low proficiency learners in the untimed task, in number. Surprisingly, both groups showed the opposite pattern in gender, performing more accurately with Default than Feature Clash errors. Thus, these learners were more accurate with (52) than (53) below:

52) La casa es \*viejo. **Gender default**  
 The<sub>femsg</sub> house<sub>femsg</sub> is old<sub>Msg</sub>

53) El libro es \*vieja. **Gender feature-clash**  
 The<sub>msg</sub> book<sub>msg</sub> is old<sub>FEMsg</sub>

This pattern wasn't expected, since it is incompatible with what has consistently been reported in the literature (e.g. McCarthy, 2008; Prévost & White, 2000; White et al., 2004; Montrul et al.,

2008), namely that learners are more accurate with feature clash than with default type errors, both in production and comprehension. One possible explanation for this inconsistency that we have explored is the existence of differences in the frequency of the adjectives used in our experiment. Since  $d'$  scores indicate sensitivity to violations in relation to grammatical counterparts, we compared the frequency of the adjectives in their ungrammatical forms (default and feature clash) to the frequency of their grammatical form in the following way:

(54)  $\text{intensos}_{\text{MSg}}$  vs.  $*\text{intensa}_{\text{FemSg}}$  (grammatical vs. gender feature clash sg.)

(55)  $\text{adictiva}_{\text{FemSg}}$  vs.  $*\text{adictivo}_{\text{MSg}}$  (grammatical vs. gender default sg.)

(56)  $\text{intensos}_{\text{MPI}}$  vs.  $*\text{intensas}_{\text{FemPI}}$  (grammatical vs. gender feature clash pl.)

(57)  $\text{adictivas}_{\text{FemPI}}$  vs.  $*\text{adictivos}_{\text{MPI}}$  (grammatical vs. gender default pl.)

However, we found no significant differences for these comparisons.

Another possible explanation for this pattern is that the nature of our task is different from the tasks used in previous studies. In these studies, learners were asked to supply an agreeing form in some way, while in our task, we provide learners with the different violations and ask them to detect them. It could be possible that in this type of task learners were more sensitive to having the marked feature (feminine) come first, making it easier for them to predict the upcoming feature later in the sentence (see Wagers et al., 2009). A possible reason for why this wouldn't hold for number could be the extra morpheme (-s) attached to the adjective in number feature clash errors, which would give them special saliency. In fact, the scores for feature clash conditions were remarkably high, even at the low levels of proficiency. However, why our learners were more accurate with gender default than gender feature clash errors is still an open question that requires further research.

It is also important to emphasize that this pattern is not predicted by the theories we are testing: Failed Functional Features Hypothesis and Interpretability Hypothesis (representational accounts) predict general variability, not a specific asymmetry in learner errors (Hawkins, 2001). Full Transfer/Full Access and Missing Surface Inflection Hypothesis (computational accounts) predict the opposite pattern, better performance with feature-clash than default errors, because of the relative computational ease to access the unmarked feature under processing burden. This, in turn, would make default errors harder to detect, contrary to what we found in the case of gender. Therefore, further research is needed to provide a more precise account for this pattern. Data collection is ongoing, and it might shed more light on this issue.

Leaving aside the pattern observed for gender, the variability that emerged in the speeded task for the intermediate group was no longer present in the Untimed GJT, which points at increased computational demands as the main factor in the results obtained in the speeded task. Again, this task effect is more consistent with computational accounts, since representational accounts predict variability *regardless of task*.

Finally, recall McCarthy's proposal that variability in low proficiency learners is due to a representational deficit in the L2 morphology. McCarthy specifically proposed that at low levels of proficiency the marked feature hasn't been acquired yet, yielding unsystematic or bidirectional errors, rather than systematic use of defaults. McCarthy didn't have a low proficiency group in her (2008) study. However, she observed this pattern in some of her intermediate learners, and argued that they could be at an initial stage acquiring the gender features. Our results for the low proficiency group seem to challenge a morphological deficit at early stages of acquisition. As we saw in the feature analysis (Gender vs. Number), the low proficiency learners showed higher accuracy with number than gender, in both the speeded and

untimed tasks, pointing at transfer effects. Crucially, in addition to this feature effect, in the untimed task they also showed higher accuracy with default than feature clash errors in gender. The presence of this asymmetry, as opposed to bidirectional errors, suggests that the marked feature (feminine) *has* in fact been acquired. Thus, our results indicate that transfer plays an important role in initial stages of acquisition of the L2 morphology. However, a morphological deficit is not supported. On the other hand, individual data shouldn't be ignored, if the aim is to provide an accurate account of the nature of morphological variability. Along with McCarthy (2008), other studies have reported variability at the individual level (Bruhn de Garavito & White, 2000; Hawkins, 1998), with some learners adopting the feminine, rather than the masculine, as a default form, and others showing no evidence of error systematicity or use of defaults. Given that there is a clear pattern of variability that accounts for the majority of learners, it could be hypothesized that the patterns observed in these other learners are related to individual differences. In order to test this hypothesis, future studies investigating morphological variability should include different measures of individual characteristics, such as working memory, verbal ability, or aptitude, among others. This approach would allow us to investigate whether specific individual traits are associated with these 'atypical' patterns observed in some learners.

## **9.2 Native Speakers**

As for the native speakers, our first question asked whether task demands would impact this group's accuracy and sensitivity to the different types of violations tested. In the Speed 2 and 3 groups we observed similar declines in accuracy to the learners' in the Speeded GJT. In fact, when we compared the native speakers in Speed 2 and the advanced learners in the Speeded

GJT, we found no significant differences in overall accuracy between the two groups. This quantitative similarity between the two populations is consistent with accounts positing the same abstract representations for learners and native speakers.

In our second question we specifically inquired about differences in the detection of the different violations we were testing. We found that native speakers showed evidence of default morphology as speed increased. Specifically, the Speed 2 group was more sensitive to Feature Clash than Default errors in number, similar to intermediate learners in the speeded task. Both groups reflect the same feature asymmetry for number when performing under increased task demands. This qualitative similarity is more consistent with computational accounts of morphological variability.

On the other hand, native speakers didn't show differences detecting feature clash versus default errors in gender. In order to explain this lack of correspondence between the patterns found for gender and number we turned to the native literature. To the best of our knowledge there are no studies investigating default morphology in native speakers of Spanish. However, there is extensive research on the native processing of gender and number agreement, with some studies pointing at differences between the processing of the two features (Faussart et al., 1999; De Vincenzi, 1999, in Italian; Di Domenico and De Vincenzi, 1995, in Italian; Vigliocco et al., 1996b; Igoa et al., 1999) and others finding no differences (Barber and Carreiras, 2003; Osterhout and Mobley, 1995, with English pronouns; Hagoort and Brown, 1999, in Dutch; Gunter et al., 2000 in German; Deutsch and Bentin, 2001, in Hebrew; Barber and Carreiras, 2005). Nevertheless, it seems that the evidence pointing at no differences between the two features is more robust (Acuña-Fariña, 2009). This conclusion is consistent with our results, since we didn't find feature effects in any of the native speaker groups, but it doesn't explain the

within feature differences we found for number, but not for gender. A possible explanation could be that performance on the grammatical conditions masked existing differences between the ungrammatical conditions. Recall that  $d'$  scores result from the relationship between acceptance and rejection of grammatical and ungrammatical conditions. Therefore, errors on the grammatical items could actually have suppressed some existing effects. This explanation could also account for the fact that no effects were observed at the highest speed (Speed 3 group). Therefore, this possibility will be explored in a future study by analyzing the present data based on acceptance rates for the ungrammatical conditions only.

## **10. Conclusion**

The present study explored the nature of variability in gender and number agreement in learners and native speakers of Spanish. Different claims about the sources of morphological variability are made by different accounts depending on the importance they concede to divergences between the feature inventories in the native and second language, as opposed to computational difficulties. We approached this issue by testing sensitivity to different types of agreement errors under different task demands. Our results suggest that in contrast to predictions by representational accounts of morphological variability, the lack of an uninterpretable feature like gender in the native language is not a limiting factor to acquiring it in a second language. Our intermediate and advanced learners in the untimed task didn't show differences in performance with gender, a feature absent from their native language, versus number, a feature shared by their native and their second language. In addition, two of the four advanced learners in this task showed perfect accuracy in every condition we were testing.

In the speeded task we did observe an advantage for number over gender, indicating that L1 transfer can be a modulating factor in computation under an additional processing load (Hopp, 2010). In this task specific patterns of variability were evident in the intermediate group, confirming that learner variability is systematic, rather than random. Crucially, these patterns didn't emerge in the untimed task, showing again the impact of task demands on learner variability.

Interestingly, the pattern observed for number in intermediate learners in the speeded task was also present for native speakers in Speed 2, showing that increased task demands had a similar impact on both learners and natives.

In conclusion, the main contributions of the present study are the following:

- a) It has provided evidence of L2 native-like performance in agreement with both gender and number in the absence of additional computational burden.
- b) It has shown how increased task demands result in agreement variability in levels of proficiency where invariable performance is observed in a regular grammaticality judgment task.
- c) It has found evidence of default morphology in native speakers performing under increased computational burden.

Taken together, although further research on the patterns observed is needed, these findings are more consistent with computational accounts of morphological variability.

## References

Acuña-Fariña, J.C. (2009). The linguistics and psycholinguistics of agreement: a tutorial overview. *Lingua*, 119, 389–424.



- Alemán-Bañón, J. (2010). The Processing of Gender and Number Agreement in Spanish: an ERP investigation. MA thesis, University of Kansas.
- Bader, M., & Meng, M. (1999b). Subject-object ambiguities in German embedded clauses: an across-the-board comparison. *Journal of Psycholinguistic Research*, 28, 121–143.
- Barber, H. & Carreiras, M. (2005). Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience*, 17, 137-153.
- Barber, H. & Carreiras, M. (2003). Integrating gender and number information in Spanish word pairs: An ERP study. *Cortex*, 39, 465–482.
- Battistella, E. (1990). *Markedness: The Evaluative Superstructure of Language*, SUNY Press, Albany.
- Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis*, 20, 3-49.
- Bley-Vroman, R. (1989). The logical problem of second language learning. In S. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp, 41-68). Cambridge: Cambridge University Press.
- Bobaljik, J. D. (2002). A-chains at the PF-interface: Copies and ‘covert’ Movement, *Natural Language and Linguistic Theory* 20, 197–267.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Carroll, S. (1989). Second-language acquisition and the computational paradigm. *Language Learning*, 39, 535-594.
- Carstens, V. (2000). Concord in minimalist theory. *Linguistic Inquiry*, 31, 319–355.
- Clahsen, H., Eisenbeiss, S., & Penke, M. (1996). Lexical learning in early syntactic development. In H. Clahsen (Ed.), *Generative Perspective on Language Acquisition* (pp. 129-160). Amsterdam: John Benjamins.
- Clahsen, H., & Muysken, P. (1986). The UG paradox in L2 acquisition. *Second Language Research*, 5, 1-29.
- Corbett, G. G. (1991). *Gender*. New York: Cambridge University Press.
- Cowper, E. (2005): The geometry of interpretable features: Infl in English and Spanish. *Language*, 81, 10–46.

- Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: evidence from ERPs and eye movements. *Journal of Memory and Language*, 45, 200–224.
- De Vincenzi, M. (1999). Differences between the morphology of gender and number: evidence from establishing coreferences. *Journal of Psycholinguistic Research* 28, 537–553.
- Di Domenico, E., & De Vincenzi, M. (1995). Gender and number in the retrieval of pronoun antecedents: differences in use and representation. In: Nash, L., Tsoules, F., Zribi-Herts, A. (Eds.), *Actes du deuxie`me colloque 'Langues et grammaire'* (pp. 95–109). Paris.
- Eisenbeiss, S. (1994). Kasus und Wortstellungsvariation im deutschen Mittelfeld. Theoretische Überlegungen und Untersuchungen zum Erstspracherwerb. In B. Haftka (Ed.), *Was Determiniert Wortstellungsvariation?* (pp. 279-298). Opladen: Westdeutscher Verlag.
- Eisenbeiss, S. (2002). *Merkmalsgesteuerter Grammatikerwerb. Eine Untersuchung zum Erwerb der Struktur und Flexion der Nominalphrase*. Unpublished Ph.D. dissertation, University of Düsseldorf, Düsseldorf.
- Eisenbeiss, S., Bartke, S., & Clahsen, H. (2005/2006). Structural and lexical case in child German: Evidence from language-impaired and typically-developing children. *Language Acquisition*, 13, 3-32.
- Faussart, C., Jacobowicz, C., & Costes, M. (1999). Gender and number processing in spoken French and Spanish. *Rivista de Linguistica*, 11, 75–101.
- Franceschina, F. (2005). *Fossilized Second Language Grammars: The Acquisition of Grammatical Gender*. Amsterdam: John Benjamins.
- Franceschina, F. (2002). Case and (phi) feature agreement in advanced L2 Spanish grammars. *EUROSLA Yearbook* (Vol. 2, pp. 71-86). Amsterdam: Benjamins.
- Gunter, T.C., Friederici, A.D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, 12, 556–568.
- Hagoort, P., & Brown, C.M. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing, *Journal of Psycholinguistic Research*, 28, 715–728.
- Halle, M. (1997). Distributed Morphology: Impoverishment and fission. In *MIT working papers in linguistics 30: PF-Papers at the interface* (pp. 425-449). MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
- Halle, M. & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S.J. Keyser (Eds.), *The view from building 20: essays in linguistics in honor of Sylvain Bromberger* (pp. 111–76). Cambridge MA: MIT Press.

- Harley, H. (1994). Hug a tree: deriving the morphosyntactic feature geometry. In Carnie, A., & Harley, H. (Eds.), *MITWPL 21: Papers on phonology and morphology* (pp. 275-288). Cambridge, Mass: Department of Linguistics and Philosophy, MIT.
- Harley, H. & Ritter, E. (2002). Structuring the bundle: A universal morphosyntactic feature geometry. In H. Simon & H. Wiese (Eds.), *Pronouns: Features and representation* (pp. 23-39). Amsterdam: John Benjamins.
- Harris, J. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, 22, 27–62.
- Hawkins, R. (2001). The theoretical significance of Universal Grammar in second language Acquisition. *Second Language Research* 21, 345–367.
- Hawkins, R. (1998). *The inaccessibility of formal features of functional categories in second language acquisition*. Paper presented at the Pacific Second Language Research Forum (PacSlrf), Tokyo.
- Hawkins, R., & Chan, C. Y.-h. (1997). The partial availability of universal grammar in second language acquisition: The “failed functional features hypothesis.” *Second Language Research*, 22, 369-397.
- Haznedar, B. & Schwartz, B. (1997). Are there optional infinitives in child L2 acquisition? In Hughes, E. Hughes, M. and Greenhill, A., editors, *Proceedings of the 21st Annual Boston University Conference on Language Development (BUCLD)* (pp. 293–306). Somerville, MA: Cascadilla Press.
- Hopp, H., (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, 120, 901-931.
- Igoa, J.M., García-Albea, J.E., & Sánchez-Casas, R. (1999). Gender-number dissociations in sentence production in Spanish. *Rivista di Linguistica*, 11, 163–196.
- Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: an eye-movement investigation. *Language Learning*, 59, 503-535.
- Lardiere, D. (2000). Mapping features to forms in second language acquisition. In J. Archibald (Ed.). *Second language acquisition and linguistic theory* (pp. 102-129). Malden, MA: Blackwell.
- McCarthy, C. (2008). Morphological variability in the comprehension of agreement: an argument for representation over computation. *Second Language Research*, 24, 459-486.
- McCarthy, C. (2007). *Morphological variability in second language Spanish*. Unpublished PhD dissertation, McGill University, Montreal.

- McDonald, J. L. (2006). Beyond the critical period: processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55, 381-401.
- Meng, M., & Bader, M., (2000). Ungrammaticality detection and garden-path strength: evidence for serial parsing. *Language and Cognitive Processes*, 6, 615-666.
- Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: the effects of age and context of acquisition. *Language Learning*, 58:3, 503-553.
- Montrul, S. (2005). Second language acquisition and first language loss in adult early bilinguals: Exploring some differences and similarities. *Second Language Research*, 21, 1-51.
- Osterhout, L., & Mobley, L.A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739-773.
- Picallo, M. C. (1991). Nominals and nominalizations in Catalan. *Probus*, 3, 279-316.
- Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16, 110-133.
- Ritter, E. (1993). Where is gender? *Linguistic Inquiry*, 24, 795-803.
- Ritter, E. (1991). Two functional categories in noun phrases: Evidence from Modern Hebrew. In S. D. Rothstein (Ed.), *Perspectives on Phrase Structure: Heads and Licensing*, *Syntax and Semantics* (Vol. 25, pp. 37-62). New York: Academic Press.
- Schlesewsky, M., & Frisch, S. (2003). Nominative case as a multidimensional default. In S. Fischer & R. van de Vijver & R. Vogel (Eds.), *Linguistics in Potsdam* (Vol. 21, pp. 61-72).
- Schachter, J. (1990). On the issue of completeness in second language acquisition. *Second Language Research*, 6, 93-124.
- Schütze, C. T. (1997). *INFL in Child and Adult language: Agreement, Case and Licensing*. Unpublished Ph.D. dissertation, MIT, Cambridge, MA.
- Schwartz, B. D., & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, 12, 40-72.
- Sorace, A. (2005). Selective optionality in language development. In L. Cornips and K. Corrigan (Eds.), *Syntax and Variation. Reconciling the Biological and the Social* (pp. 55-80). Amsterdam: John Benjamins.

- Tagliaferri, B. (2005). *Paradigm*. Perception Research Systems, Inc.  
www.perceptionresearchsystems.com
- Tsimpli, I., & Dimitrakopoulou, M. (2007). The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, 23, 215-242.
- Vigliocco, G., Butterworth, B., & Garrett, M.F. (1996b). Subject–verb agreement in Spanish and English: the role of conceptual factors. *Cognition*, 51, 261–298.
- Vogel, R., & Frisch, S. (2003). The Resolution of Case Conflicts – a Pilot Study.  
*Linguistics in Potsdam*, 21, 91–103.
- Wagers, M., Lau, E., & Phillips C. (2009). “Encoding syntactic predictions: evidence from the dynamics of agreement.” CUNY Sentence Processing Conference. Davis, CA. 26 March, 2009.
- White, L., Valenzuela, E., Kozłowska-MacGregor, M., Leung, I. Y.-K., & Ben Ayed, H. (2001). The status of abstract features in interlanguage: Gender and number in L2 Spanish. In A. H. J. Do, L. Dominguez, & A. Johansen (Eds.), *BUCLD 25: Proceedings of the 25<sup>th</sup> annual Boston University Conference on Language Development* (2000) (pp. 792-802). Somerville, MA: Cascadia.
- White, L., Valenzuela, E., Kozłowska-MacGregor, M., & Leung, Y.-K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, 25, 105-133.

## APPENDIX. Target stimuli

Feminine nouns:

1. Juan dijo que vio una película que era romántica en París.  
Juan dijo que vio una película que era romántico en París.  
Juan dijo que vio una película que era románticas en París.  
Juan dijo que vio unas películas que eran románticas en París.  
Juan dijo que vio unas películas que eran romántica en París.  
Juan dijo que vio unas películas que eran románticos en París.  
“Juan said that he saw a/some movie(s) that was/were romantic in Paris.”
2. Juan dijo que comió una salsa que era adictiva en Gibraltar.  
Juan dijo que comió una salsa que era adictivo en Gibraltar.  
Juan dijo que comió una salsa que era adictivas en Gibraltar.  
Juan dijo que comió unas salsas que eran adictivas en Gibraltar.  
Juan dijo que comió unas salsas que eran adictiva en Gibraltar.  
Juan dijo que comió unas salsas que eran adictivos en Gibraltar.
3. Esther dijo que compró una falda que era fresca en Milán.  
Esther dijo que compró una falda que era fresco en Milán.  
Esther dijo que compró una falda que era frescas en Milán.  
Esther dijo que compró unas faldas que eran frescas en Milán.  
Esther dijo que compró unas faldas que eran fresca en Milán.  
Esther dijo que compró unas faldas que eran frescos en Milán.
4. Esther dijo que visitó una plaza que era inmensa en Roma.  
Esther dijo que visitó una plaza que era inmenso en Roma.  
Esther dijo que visitó una plaza que era inmensas en Roma.  
Esther dijo que visitó unas plazas que eran inmensas en Roma.  
Esther dijo que visitó unas plazas que eran inmensa en Roma.  
Esther dijo que visitó unas plazas que eran inmensos en Roma.
5. Manuel dijo que encontró una mesa que era pesada en Australia.  
Manuel dijo que encontró una mesa que era pesado en Australia.  
Manuel dijo que encontró una mesa que era pesadas en Australia.  
Manuel dijo que encontró unas mesas que eran pesadas en Australia.  
Manuel dijo que encontró unas mesas que eran pesada en Australia.  
Manuel dijo que encontró unas mesas que eran pesados en Australia.
6. Manuel dijo que compró una espada que era afilada en Edimburgo.  
Manuel dijo que compró una espada que era afilado en Edimburgo.  
Manuel dijo que compró una espada que era afiladas en Edimburgo.  
Manuel dijo que compró unas espadas que eran afiladas en Edimburgo.  
Manuel dijo que compró unas espadas que eran afilada en Edimburgo.  
Manuel dijo que compró unas espadas que eran afilados en Edimburgo.

7. Laura dijo que encontró una pastelería que era minúscula en Italia.  
Laura dijo que encontró una pastelería que era minúsculo en Italia.  
Laura dijo que encontró una pastelería que era minúsculas en Italia.  
Laura dijo que encontró unas pastelerías que eran minúsculas en Italia.  
Laura dijo que encontró unas pastelerías que eran minúscula en Italia.  
Laura dijo que encontró unas pastelerías que eran minúsculos en Italia.
8. Laura dijo que visitó una academia que era prestigiosa en Inglaterra.  
Laura dijo que visitó una academia que era prestigioso en Inglaterra.  
Laura dijo que visitó una academia que era prestigiosas en Inglaterra.  
Laura dijo que visitó unas academias que eran prestigiosas en Inglaterra.  
Laura dijo que visitó unas academias que eran prestigiosa en Inglaterra.  
Laura dijo que visitó unas academias que eran prestigiosos en Inglaterra.
9. Miguel dijo que compró una corona que era auténtica en Londres.  
Miguel dijo que compró una corona que era auténtico en Londres.  
Miguel dijo que compró una corona que era auténticas en Londres.  
Miguel dijo que compró unas coronas que eran auténticas en Londres.  
Miguel dijo que compró unas coronas que eran auténtica en Londres.  
Miguel dijo que compró unas coronas que eran auténticos en Londres.
10. Miguel dijo que compró una guitarra que era bonita en Miami.  
Miguel dijo que compró una guitarra que era bonito en Miami.  
Miguel dijo que compró una guitarra que era bonitas en Miami.  
Miguel dijo que compró unas guitarras que eran bonitas en Miami.  
Miguel dijo que compró unas guitarras que eran bonita en Miami.  
Miguel dijo que compró unas guitarras que eran bonitos en Miami.
11. María dijo que visitó una biblioteca que era gratuita en Portugal.  
María dijo que visitó una biblioteca que era gratuito en Portugal.  
María dijo que visitó una biblioteca que era gratuitas en Portugal.  
María dijo que visitó unas bibliotecas que eran gratuitas en Portugal.  
María dijo que visitó unas bibliotecas que eran gratuita en Portugal.  
María dijo que visitó unas bibliotecas que eran gratuitos en Portugal.
12. María dijo que vio una montaña que era rocosa en Colorado.  
María dijo que vio una montaña que era rocoso en Colorado.  
María dijo que vio una montaña que era rocosas en Colorado.  
María dijo que vio unas montañas que eran rocosas en Colorado.  
María dijo que vio unas montañas que eran rocosa en Colorado.  
María dijo que vio unas montañas que eran rocosos en Colorado.

13. Inés dijo que vio una sierra que era grandiosa en España.  
Inés dijo que vio una sierra que era grandioso en España.  
Inés dijo que vio una sierra que era grandiosas en España.  
Inés dijo que vio unas sierras que eran grandiosas en España.  
Inés dijo que vio unas sierras que eran grandiosa en España.  
Inés dijo que vio unas sierras que eran grandiosos en España.
14. Inés dijo que comió una fresa que era redonda en Irlanda.  
Inés dijo que comió una fresa que era redondo en Irlanda.  
Inés dijo que comió una fresa que era redondas en Irlanda.  
Inés dijo que comió unas fresas que eran redondas en Irlanda.  
Inés dijo que comió unas fresas que eran redonda en Irlanda.  
Inés dijo que comió unas fresas que eran redondos en Irlanda.
15. Pedro dijo que visitó una bodega que era famosa en Francia.  
Pedro dijo que visitó una bodega que era famoso en Francia.  
Pedro dijo que visitó una bodega que era famosas en Francia.  
Pedro dijo que visitó unas bodegas que eran famosas en Francia.  
Pedro dijo que visitó unas bodegas que eran famosa en Francia.  
Pedro dijo que visitó unas bodegas que eran famosos en Francia.
16. Pedro dijo que compró una corbata que era azulada en Cannes.  
Pedro dijo que compró una corbata que era azulado en Cannes.  
Pedro dijo que compró una corbata que era azuladas en Cannes.  
Pedro dijo que compró unas corbatas que eran azuladas en Cannes.  
Pedro dijo que compró unas corbatas que eran azulada en Cannes.  
Pedro dijo que compró unas corbatas que eran azulados en Cannes.
17. Pablo dijo que encontró una isla que era preciosa en Europa.  
Pablo dijo que encontró una isla que era precioso en Europa.  
Pablo dijo que encontró una isla que era preciosas en Europa.  
Pablo dijo que encontró unas islas que eran preciosas en Europa.  
Pablo dijo que encontró unas islas que eran preciosa en Europa.  
Pablo dijo que encontró unas islas que eran preciosos en Europa.
18. Pablo dijo que vio una península que era gigantesca en Europa.  
Pablo dijo que vio una península que era gigantesco en Europa.  
Pablo dijo que vio una península que era gigantescas en Europa.  
Pablo dijo que vio unas penínsulas que eran gigantescas en Europa.  
Pablo dijo que vio unas penínsulas que eran gigantesca en Europa.  
Pablo dijo que vio unas penínsulas que eran gigantescos en Europa.



19. Alicia dijo que comió una lechuga que era nutritiva en Tailandia.  
Alicia dijo que comió una lechuga que era nutritivo en Tailandia.  
Alicia dijo que comió una lechuga que era nutritivas en Tailandia.  
Alicia dijo que comió unas lechugas que eran nutritivas en Tailandia.  
Alicia dijo que comió unas lechugas que eran nutritiva en Tailandia.  
Alicia dijo que comió unas lechugas que eran nutritivos en Tailandia.
20. Alicia dijo que comió una naranja que era jugosa en California.  
Alicia dijo que comió una naranja que era jugoso en California.  
Alicia dijo que comió una naranja que era jugosas en California.  
Alicia dijo que comió unas naranjas que eran jugosas en California.  
Alicia dijo que comió unas naranjas que eran jugosa en California.  
Alicia dijo que comió unas naranjas que eran jugosos en California.
21. José dijo que comió una ensalada que era insípida en Moscú.  
José dijo que comió una ensalada que era insípido en Moscú.  
José dijo que comió una ensalada que era insípidas en Moscú.  
José dijo que comió unas ensaladas que eran insípidas en Moscú.  
José dijo que comió unas ensaladas que eran insípida en Moscú.  
José dijo que comió unas ensaladas que eran insípidos en Moscú.
22. José dijo que visitó una zona que era misteriosa en Brasil.  
José dijo que visitó una zona que era misterioso en Brasil.  
José dijo que visitó una zona que era misteriosas en Brasil.  
José dijo que visitó unas zonas que eran misteriosas en Brasil.  
José dijo que visitó unas zonas que eran misteriosa en Brasil.  
José dijo que visitó unas zonas que eran misteriosos en Brasil.
23. Marta dijo que vio una costa que era monótona en Hawai.  
Marta dijo que vio una costa que era monótono en Hawai.  
Marta dijo que vio una costa que era monótonas en Hawai.  
Marta dijo que vio unas costas que eran monótonas en Hawai.  
Marta dijo que vio unas costas que eran monótona en Hawai.  
Marta dijo que vio unas costas que eran monótonos en Hawai.
24. Marta dijo que compró una esmeralda que era valiosa en Sudáfrica.  
Marta dijo que compró una esmeralda que era valioso en Sudáfrica.  
Marta dijo que compró una esmeralda que era valiosas en Sudáfrica.  
Marta dijo que compró unas esmeraldas que eran valiosas en Sudáfrica.  
Marta dijo que compró unas esmeraldas que eran valiosa en Sudáfrica.  
Marta dijo que compró unas esmeraldas que eran valiosos en Sudáfrica.

25. Ramón dijo que encontró una turquesa que era hermosa en Egipto.  
Ramón dijo que encontró una turquesa que era hermoso en Egipto.  
Ramón dijo que encontró una turquesa que era hermosas en Egipto.  
Ramón dijo que encontró unas turquesas que eran hermosas en Egipto.  
Ramón dijo que encontró unas turquesas que eran hermosa en Egipto.  
Ramón dijo que encontró unas turquesas que eran hermosos en Egipto.
26. Ramón dijo que encontró una hamburguesa que era apetitosa en Chicago.  
Ramón dijo que encontró una hamburguesa que era apetitoso en Chicago.  
Ramón dijo que encontró una hamburguesa que era apetitosas en Chicago.  
Ramón dijo que encontró unas hamburguesas que eran apetitosas en Chicago.  
Ramón dijo que encontró unas hamburguesas que eran apetitosa en Chicago.  
Ramón dijo que encontró unas hamburguesas que eran apetitosos en Chicago.
27. Ana dijo que compró una cerveza que era aromática en Berlín.  
Ana dijo que compró una cerveza que era aromático en Berlín.  
Ana dijo que compró una cerveza que era aromáticas en Berlín.  
Ana dijo que compró unas cervezas que eran aromáticas en Berlín.  
Ana dijo que compró unas cervezas que eran aromática en Berlín.  
Ana dijo que compró unas cervezas que eran aromáticos en Berlín.
28. Ana dijo que vio una tela que era fina en París.  
Ana dijo que vio una tela que era fino en París.  
Ana dijo que vio una tela que era finas en París.  
Ana dijo que vio unas telas que eran finas en París.  
Ana dijo que vio unas telas que eran fina en París.  
Ana dijo que vio unas telas que eran finos en París.
29. Esteban dijo que compró una camiseta que era cómoda en Roma.  
Esteban dijo que compró una camiseta que era cómodo en Roma.  
Esteban dijo que compró una camiseta que era cómodas en Roma.  
Esteban dijo que compró unas camisetas que eran cómodas en Roma.  
Esteban dijo que compró unas camisetas que eran cómoda en Roma.  
Esteban dijo que compró unas camisetas que eran cómodos en Roma.
30. Esteban dijo que encontró una chaqueta que era cálida en Milán.  
Esteban dijo que encontró una chaqueta que era cálido en Milán.  
Esteban dijo que encontró una chaqueta que era cálidas en Milán.  
Esteban dijo que encontró unas chaquetas que eran cálidas en Milán.  
Esteban dijo que encontró unas chaquetas que eran cálida en Milán.  
Esteban dijo que encontró unas chaquetas que eran cálidos en Milán.

31. Alejandra dijo que comió una frambuesa que era curativa en Ecuador.  
Alejandra dijo que comió una frambuesa que era curativo en Ecuador.  
Alejandra dijo que comió una frambuesa que era curativas en Ecuador.  
Alejandra dijo que comió unas frambuesas que eran curativas en Ecuador.  
Alejandra dijo que comió unas frambuesas que eran curativa en Ecuador.  
Alejandra dijo que comió unas frambuesas que eran curativos en Ecuador.
32. Alejandra dijo que encontró una perla que era delicada en Italia.  
Alejandra dijo que encontró una perla que era delicado en Italia.  
Alejandra dijo que encontró una perla que era delicadas en Italia.  
Alejandra dijo que encontró unas perlas que eran delicadas en Italia.  
Alejandra dijo que encontró unas perlas que eran delicada en Italia.  
Alejandra dijo que encontró unas perlas que eran delicados en Italia.
33. Jorge dijo que comió una sopa que era amarilla en Bolivia.  
Jorge dijo que comió una sopa que era amarillo en Bolivia.  
Jorge dijo que comió una sopa que era amarillas en Bolivia.  
Jorge dijo que comió unas sopas que eran amarillas en Bolivia.  
Jorge dijo que comió unas sopas que eran amarilla en Bolivia.  
Jorge dijo que comió unas sopas que eran amarillos en Bolivia.
34. Jorge dijo que vio una danza que era complicada en África.  
Jorge dijo que vio una danza que era complicado en África.  
Jorge dijo que vio una danza que era complicadas en África.  
Jorge dijo que vio unas danzas que eran complicadas en África.  
Jorge dijo que vio unas danzas que eran complicada en África.  
Jorge dijo que vio unas danzas que eran complicados en África.
35. Gloria dijo que encontró una lavadora que era metálica en Japón.  
Gloria dijo que encontró una lavadora que era metálico en Japón.  
Gloria dijo que encontró una lavadora que era metálicas en Japón.  
Gloria dijo que encontró unas lavadoras que eran metálicas en Japón.  
Gloria dijo que encontró unas lavadoras que eran metálica en Japón.  
Gloria dijo que encontró unas lavadoras que eran metálicos en Japón.
36. Gloria dijo que visitó una sauna que era amplia en Grecia.  
Gloria dijo que visitó una sauna que era amplio en Grecia.  
Gloria dijo que visitó una sauna que era amplias en Grecia.  
Gloria dijo que visitó unas saunas que eran amplias en Grecia.  
Gloria dijo que visitó unas saunas que eran amplia en Grecia.  
Gloria dijo que visitó unas saunas que eran amplios en Grecia.

37. Germán dijo que vio una boda que era tranquila en Noruega.  
Germán dijo que vio una boda que era tranquilo en Noruega.  
Germán dijo que vio una boda que era tranquilas en Noruega.  
Germán dijo que vio unas bodas que eran tranquilas en Noruega.  
Germán dijo que vio unas bodas que eran tranquila en Noruega.  
Germán dijo que vio unas bodas que eran tranquilos en Noruega.
38. Germán dijo que comió una pizza que era digestiva en Nápoles.  
Germán dijo que comió una pizza que era digestivo en Nápoles.  
Germán dijo que comió una pizza que era digestivas en Nápoles.  
Germán dijo que comió unas pizzas que eran digestivas en Nápoles.  
Germán dijo que comió unas pizzas que eran digestiva en Nápoles.  
Germán dijo que comió unas pizzas que eran digestivos en Nápoles.
39. Fran dijo que visitó una iglesia que era moderna en Liverpool.  
Fran dijo que visitó una iglesia que era moderno en Liverpool.  
Fran dijo que visitó una iglesia que era modernas en Liverpool.  
Fran dijo que visitó unas iglesias que eran modernas en Liverpool.  
Fran dijo que visitó unas iglesias que eran moderna en Liverpool.  
Fran dijo que visitó unas iglesias que eran modernos en Liverpool.
40. Fran dijo que encontró una huella que era humana en Atapuerca.  
Fran dijo que encontró una huella que era humano en Atapuerca.  
Fran dijo que encontró una huella que era humanas en Atapuerca.  
Fran dijo que encontró unas huellas que eran humanas en Atapuerca.  
Fran dijo que encontró unas huellas que eran humana en Atapuerca.  
Fran dijo que encontró unas huellas que eran humanos en Atapuerca.
41. Elvira dijo que compró una cámara que era precisa en Tokio.  
Elvira dijo que compró una cámara que era preciso en Tokio.  
Elvira dijo que compró una cámara que era precisas en Tokio.  
Elvira dijo que compró unas cámaras que eran precisas en Tokio.  
Elvira dijo que compró unas cámaras que eran precisa en Tokio.  
Elvira dijo que compró unas cámaras que eran precisos en Tokio.
42. Elvira dijo que vio una impresora que era compleja en China.  
Elvira dijo que vio una impresora que era complejo en China.  
Elvira dijo que vio una impresora que era complejas en China.  
Elvira dijo que vio unas impresoras que eran complejas en China.  
Elvira dijo que vio unas impresoras que eran compleja en China.  
Elvira dijo que vio unas impresoras que eran complejos en China.

43. Carolina dijo que comió una manzana que era alimenticia en París.  
Carolina dijo que comió una manzana que era alimenticio en París.  
Carolina dijo que comió una manzana que era alimenticias en París.  
Carolina dijo que comió unas manzanas que eran alimenticias en París.  
Carolina dijo que comió unas manzanas que eran alimenticia en París.  
Carolina dijo que comió unas manzanas que eran alimenticios en París.
44. Carolina dijo que visitó una selva que era húmeda en África.  
Carolina dijo que visitó una selva que era húmedo en África.  
Carolina dijo que visitó una selva que era húmedas en África.  
Carolina dijo que visitó unas selvas que eran húmedas en África.  
Carolina dijo que visitó unas selvas que eran húmeda en África.  
Carolina dijo que visitó unas selvas que eran húmedos en África.
45. Arturo dijo que visitó una provincia que era rica en Senegal.  
Arturo dijo que visitó una provincia que era rico en Senegal.  
Arturo dijo que visitó una provincia que era ricas en Senegal.  
Arturo dijo que visitó unas provincias que eran ricas en Senegal.  
Arturo dijo que visitó unas provincias que eran rica en Senegal.  
Arturo dijo que visitó unas provincias que eran ricos en Senegal.
46. Rosario dijo que descubrió una bicicleta que era femenina en Suiza.  
Rosario dijo que descubrió una bicicleta que era femenino en Suiza.  
Rosario dijo que descubrió una bicicleta que era femeninas en Suiza.  
Rosario dijo que descubrió unas bicicletas que eran femeninas en Suiza.  
Rosario dijo que descubrió unas bicicletas que eran femenina en Suiza.  
Rosario dijo que descubrió unas bicicletas que eran femeninos en Suiza.
47. Rosario dijo que descubrió una lasaña que era sólida en Nápoles.  
Rosario dijo que descubrió una lasaña que era sólido en Nápoles.  
Rosario dijo que descubrió una lasaña que era sólidas en Nápoles.  
Rosario dijo que descubrió unas lasañas que eran sólidas en Nápoles.  
Rosario dijo que descubrió unas lasañas que eran sólida en Nápoles.  
Rosario dijo que descubrió unas lasañas que eran sólidos en Nápoles.
48. Elena dijo que descubrió una lámpara que era creativa en Madrid.  
Elena dijo que descubrió una lámpara que era creativo en Madrid.  
Elena dijo que descubrió una lámpara que era creativas en Madrid.  
Elena dijo que descubrió unas lámparas que eran creativas en Madrid.  
Elena dijo que descubrió unas lámparas que eran creativa en Madrid.  
Elena dijo que descubrió unas lámparas que eran creativos en Madrid.

49. Elena dijo que descubrió una paella que era elaborada en Turquía.  
Elena dijo que descubrió una paella que era elaborado en Turquía.  
Elena dijo que descubrió una paella que era elaboradas en Turquía.  
Elena dijo que descubrió unas paellas que eran elaboradas en Turquía.  
Elena dijo que descubrió unas paellas que eran elaborada en Turquía.  
Elena dijo que descubrió unas paellas que eran elaborados en Turquía.
50. David dijo que descubrió una verdura que era beneficiosa en Pequín.  
David dijo que descubrió una verdura que era beneficioso en Pequín.  
David dijo que descubrió una verdura que era beneficiosas en Pequín.  
David dijo que descubrió unas verduras que eran beneficiosas en Pequín.  
David dijo que descubrió unas verduras que eran beneficiosa en Pequín.  
David dijo que descubrió unas verduras que eran beneficiosos en Pequín.
51. David dijo que descubrió una pintura que era entretenida en Manhattan.  
David dijo que descubrió una pintura que era entretenido en Manhattan.  
David dijo que descubrió una pintura que era entretenidas en Manhattan.  
David dijo que descubrió unas pinturas que eran entretenidas en Manhattan.  
David dijo que descubrió unas pinturas que eran entretenida en Manhattan.  
David dijo que descubrió unas pinturas que eran entretenidos en Manhattan.
52. Ricardo dijo que descubrió una escultura que era pedagógica en Atenas.  
Ricardo dijo que descubrió una escultura que era pedagógico en Atenas.  
Ricardo dijo que descubrió una escultura que era pedagógicas en Atenas.  
Ricardo dijo que descubrió unas esculturas que eran pedagógicas en Atenas.  
Ricardo dijo que descubrió unas esculturas que eran pedagógica en Atenas.  
Ricardo dijo que descubrió unas esculturas que eran pedagógicos en Atenas.
53. Ricardo dijo que descubrió una avioneta que era sencilla en Moscú.  
Ricardo dijo que descubrió una avioneta que era sencillo en Moscú.  
Ricardo dijo que descubrió una avioneta que era sencillas en Moscú.  
Ricardo dijo que descubrió unas avionetas que eran sencillas en Moscú.  
Ricardo dijo que descubrió unas avionetas que eran sencilla en Moscú.  
Ricardo dijo que descubrió unas avionetas que eran sencillos en Moscú.
54. Marcos dijo que descubrió una lana que era fría en Brasil.  
Marcos dijo que descubrió una lana que era frío en Brasil.  
Marcos dijo que descubrió una lana que era frías en Brasil.  
Marcos dijo que descubrió unas lanas que eran frías en Brasil.  
Marcos dijo que descubrió unas lanas que eran fría en Brasil.  
Marcos dijo que descubrió unas lanas que eran fríos en Brasil.

Masculine nouns:

55. Marcos dijo que descubrió un zumo que era intenso en Cuba.  
Marcos dijo que descubrió un zumo que era intensa en Cuba.  
Marcos dijo que descubrió un zumo que era intensos en Cuba.  
Marcos dijo que descubrió unos zumos que eran intensos en Cuba.  
Marcos dijo que descubrió unos zumos que eran intenso en Cuba.  
Marcos dijo que descubrió unos zumos que eran intensas en Cuba.
56. Julio dijo que visitó un colegio que era antiguo en Londres.  
Julio dijo que visitó un colegio que era antigua en Londres.  
Julio dijo que visitó un colegio que era antiguos en Londres.  
Julio dijo que visitó unos colegios que eran antiguos en Londres.  
Julio dijo que visitó unos colegios que eran antiguo en Londres.  
Julio dijo que visitó unos colegios que eran antiguas en Londres.  
“Julio said that he visited a/some school(s) that were old in London.”
57. Julio dijo que visitó un laboratorio que era conocido en España.  
Julio dijo que visitó un laboratorio que era conocida en España.  
Julio dijo que visitó un laboratorio que era conocidos en España.  
Julio dijo que visitó unos laboratorios que eran conocidos en España.  
Julio dijo que visitó unos laboratorios que eran conocido en España.  
Julio dijo que visitó unos laboratorios que eran conocidas en España.
58. Miriam dijo que comió un pescado que era peligroso en Japón.  
Miriam dijo que comió un pescado que era peligrosa en Japón.  
Miriam dijo que comió un pescado que era peligrosos en Japón.  
Miriam dijo que comió unos pescados que eran peligrosos en Japón.  
Miriam dijo que comió unos pescados que eran peligroso en Japón.  
Miriam dijo que comió unos pescados que eran peligrosas en Japón.
59. Miriam dijo que compró un cuadro que era expresivo en Francia.  
Miriam dijo que compró un cuadro que era expresiva en Francia.  
Miriam dijo que compró un cuadro que era expresivos en Francia.  
Miriam dijo que compró unos cuadros que eran expresivos en Francia.  
Miriam dijo que compró unos cuadros que eran expresivo en Francia.  
Miriam dijo que compró unos cuadros que eran expresivas en Francia.
60. Raúl dijo que compró un gramófono que era automático en Madrid.  
Raúl dijo que compró un gramófono que era automática en Madrid.  
Raúl dijo que compró un gramófono que era automáticos en Madrid.  
Raúl dijo que compró unos gramófonos que eran automáticos en Madrid.  
Raúl dijo que compró unos gramófonos que eran automático en Madrid.  
Raúl dijo que compró unos gramófonos que eran automáticas en Madrid.

61. Raúl dijo que encontró un lago que era feo en Irlanda.  
Raúl dijo que encontró un lago que era fea en Irlanda.  
Raúl dijo que encontró un lago que era feos en Irlanda.  
Raúl dijo que encontró unos lagos que eran feos en Irlanda.  
Raúl dijo que encontró unos lagos que eran feo en Irlanda.  
Raúl dijo que encontró unos lagos que eran feas en Irlanda.
62. Mónica dijo que compró un cuchillo que era dorado en Persia.  
Mónica dijo que compró un cuchillo que era dorada en Persia.  
Mónica dijo que compró un cuchillo que era dorados en Persia.  
Mónica dijo que compró unos cuchillos que eran dorados en Persia.  
Mónica dijo que compró unos cuchillos que eran dorado en Persia.  
Mónica dijo que compró unos cuchillos que eran doradas en Persia.
63. Mónica dijo que compró un abrigo que era caluroso en Minnesota.  
Mónica dijo que compró un abrigo que era calurosa en Minnesota.  
Mónica dijo que compró un abrigo que era calurosos en Minnesota.  
Mónica dijo que compró unos abrigos que eran calurosos en Minnesota.  
Mónica dijo que compró unos abrigos que eran caluroso en Minnesota.  
Mónica dijo que compró unos abrigos que eran calurosas en Minnesota.
64. Sergio dijo que encontró un arroyo que era largo en México.  
Sergio dijo que encontró un arroyo que era larga en México.  
Sergio dijo que encontró un arroyo que era largos en México.  
Sergio dijo que encontró unos arroyos que eran largos en México.  
Sergio dijo que encontró unos arroyos que eran largo en México.  
Sergio dijo que encontró unos arroyos que eran largas en México.
65. Sergio dijo que encontró un charco que era profundo en Perú.  
Sergio dijo que encontró un charco que era profunda en Perú.  
Sergio dijo que encontró un charco que era profundos en Perú.  
Sergio dijo que encontró unos charcos que eran profundos en Perú.  
Sergio dijo que encontró unos charcos que eran profundo en Perú.  
Sergio dijo que encontró unos charcos que eran profundas en Perú.
66. Rebeca dijo que visitó un instituto que era pequeño en Portugal.  
Rebeca dijo que visitó un instituto que era pequeña en Portugal.  
Rebeca dijo que visitó un instituto que era pequeños en Portugal.  
Rebeca dijo que visitó unos institutos que eran pequeños en Portugal.  
Rebeca dijo que visitó unos institutos que eran pequeño en Portugal.  
Rebeca dijo que visitó unos institutos que eran pequeñas en Portugal.



67. Rebeca dijo que visitó un faro que era emblemático en Alejandría.  
Rebeca dijo que visitó un faro que era emblemática en Alejandría.  
Rebeca dijo que visitó un faro que era emblemáticos en Alejandría.  
Rebeca dijo que visitó unos faros que eran emblemáticos en Alejandría.  
Rebeca dijo que visitó unos faros que eran emblemático en Alejandría.  
Rebeca dijo que visitó unos faros que eran emblemáticas en Alejandría.
68. Alfonso dijo que vio un baño que era luminoso en Venecia.  
Alfonso dijo que vio un baño que era luminosa en Venecia.  
Alfonso dijo que vio un baño que era luminosos en Venecia.  
Alfonso dijo que vio unos baños que eran luminosos en Venecia.  
Alfonso dijo que vio unos baños que eran luminoso en Venecia.  
Alfonso dijo que vio unos baños que eran luminosas en Venecia.
69. Alfonso dijo que compró un chaleco que era clásico en París.  
Alfonso dijo que compró un chaleco que era clásica en París.  
Alfonso dijo que compró un chaleco que era clásicos en París.  
Alfonso dijo que compró unos chalecos que eran clásicos en París.  
Alfonso dijo que compró unos chalecos que eran clásico en París.  
Alfonso dijo que compró unos chalecos que eran clásicas en París.
70. Gonzalo dijo que vio un dormitorio que era oscuro en Dublín.  
Gonzalo dijo que vio un dormitorio que era oscura en Dublín.  
Gonzalo dijo que vio un dormitorio que era oscuros en Dublín.  
Gonzalo dijo que vio unos dormitorios que eran oscuros en Dublín.  
Gonzalo dijo que vio unos dormitorios que eran oscuro en Dublín.  
Gonzalo dijo que vio unos dormitorios que eran oscuras en Dublín.
71. Gonzalo dijo que visitó un monasterio que era ruidoso en Escocia.  
Gonzalo dijo que visitó un monasterio que era ruidosa en Escocia.  
Gonzalo dijo que visitó un monasterio que era ruidosos en Escocia.  
Gonzalo dijo que visitó unos monasterios que eran ruidosos en Escocia.  
Gonzalo dijo que visitó unos monasterios que eran ruidoso en Escocia.  
Gonzalo dijo que visitó unos monasterios que eran ruidosas en Escocia.
72. Alba dijo que encontró un camino que era florido en Francia.  
Alba dijo que encontró un camino que era florida en Francia.  
Alba dijo que encontró un camino que era floridos en Francia.  
Alba dijo que encontró unos caminos que eran floridos en Francia.  
Alba dijo que encontró unos caminos que eran florido en Francia.  
Alba dijo que encontró unos caminos que eran floridas en Francia.

73. Alba dijo que comió un pepino que era sabroso en Barcelona.  
Alba dijo que comió un pepino que era sabrosa en Barcelona.  
Alba dijo que comió un pepino que era sabrosos en Barcelona.  
Alba dijo que comió unos pepinos que eran sabrosos en Barcelona.  
Alba dijo que comió unos pepinos que eran sabroso en Barcelona.  
Alba dijo que comió unos pepinos que eran sabrosas en Barcelona.
74. Pilar dijo que visitó un cementerio que era sombrío en Holanda.  
Pilar dijo que visitó un cementerio que era sombría en Holanda.  
Pilar dijo que visitó un cementerio que era sombríos en Holanda.  
Pilar dijo que visitó unos cementerios que eran sombríos en Holanda.  
Pilar dijo que visitó unos cementerios que eran sombrío en Holanda.  
Pilar dijo que visitó unos cementerios que eran sombrías en Holanda.
75. Pilar dijo que vio un submarino que era público en Japón.  
Pilar dijo que vio un submarino que era pública en Japón.  
Pilar dijo que vio un submarino que era públicos en Japón.  
Pilar dijo que vio unos submarinos que eran públicos en Japón.  
Pilar dijo que vio unos submarinos que eran público en Japón.  
Pilar dijo que vio unos submarinos que eran públicas en Japón.
76. Álvaro dijo que compró un microscopio que era caro en Japón.  
Álvaro dijo que compró un microscopio que era cara en Japón.  
Álvaro dijo que compró un microscopio que era caros en Japón.  
Álvaro dijo que compró unos microscopios que eran caros en Japón.  
Álvaro dijo que compró unos microscopios que eran caro en Japón.  
Álvaro dijo que compró unos microscopios que eran caras en Japón.
77. Álvaro dijo que encontró un termómetro que era práctico en Korea.  
Álvaro dijo que encontró un termómetro que era práctica en Korea.  
Álvaro dijo que encontró un termómetro que era prácticos en Korea.  
Álvaro dijo que encontró unos termómetros que eran prácticos en Korea.  
Álvaro dijo que encontró unos termómetros que eran práctico en Korea.  
Álvaro dijo que encontró unos termómetros que eran prácticas en Korea.
78. Mila dijo que visitó un santuario que era simbólico en Roma.  
Mila dijo que visitó un santuario que era simbólica en Roma.  
Mila dijo que visitó un santuario que era simbólicos en Roma.  
Mila dijo que visitó unos santuarios que eran simbólicos en Roma.  
Mila dijo que visitó unos santuarios que eran simbólico en Roma.  
Mila dijo que visitó unos santuarios que eran simbólicas en Roma.

79. Mila dijo que vio un aeropuerto que era silencioso en Moscú.  
Mila dijo que vio un aeropuerto que era silenciosa en Moscú.  
Mila dijo que vio un aeropuerto que era silenciosos en Moscú.  
Mila dijo que vio unos aeropuertos que eran silenciosos en Moscú.  
Mila dijo que vio unos aeropuertos que eran silencioso en Moscú.  
Mila dijo que vio unos aeropuertos que eran silenciosas en Moscú.
80. Alberto dijo que encontró un reformatorio que era seguro en Virginia.  
Alberto dijo que encontró un reformatorio que era segura en Virginia.  
Alberto dijo que encontró un reformatorio que era seguros en Virginia.  
Alberto dijo que encontró unos reformatorios que eran seguros en Virginia.  
Alberto dijo que encontró unos reformatorios que eran seguro en Virginia.  
Alberto dijo que encontró unos reformatorios que eran seguras en Virginia.
81. Alberto dijo que comió un parmesano que era ligero en China.  
Alberto dijo que comió un parmesano que era ligera en China.  
Alberto dijo que comió un parmesano que era ligeros en China.  
Alberto dijo que comió unos parmesanos que eran ligeros en China.  
Alberto dijo que comió unos parmesanos que eran ligero en China.  
Alberto dijo que comió unos parmesanos que eran ligeras en China.
82. Raquel dijo que compró un vaso que era barato en París.  
Raquel dijo que compró un vaso que era barata en París.  
Raquel dijo que compró un vaso que era baratos en París.  
Raquel dijo que compró unos vasos que eran baratos en París.  
Raquel dijo que compró unos vasos que eran barato en París.  
Raquel dijo que compró unos vasos que eran baratas en París.
83. Raquel dijo que comió un cacao que era amargo en Sudán.  
Raquel dijo que comió un cacao que era amarga en Sudán.  
Raquel dijo que comió un cacao que era amargos en Sudán.  
Raquel dijo que comió unos cacaos que eran amargos en Sudán.  
Raquel dijo que comió unos cacaos que eran amargo en Sudán.  
Raquel dijo que comió unos cacaos que eran amargas en Sudán.
84. Roberto dijo que vio un frigorífico que era espacioso en Rusia.  
Roberto dijo que vio un frigorífico que era espaciosa en Rusia.  
Roberto dijo que vio un frigorífico que era espaciosos en Rusia.  
Roberto dijo que vio unos frigoríficos que eran espaciosos en Rusia.  
Roberto dijo que vio unos frigoríficos que eran espacioso en Rusia.  
Roberto dijo que vio unos frigoríficos que eran espaciosas en Rusia.

85. Roberto dijo que compró un violonchelo que era sobrio en Viena.  
Roberto dijo que compró un violonchelo que era sobria en Viena.  
Roberto dijo que compró un violonchelo que era sobrios en Viena.  
Roberto dijo que compró unos violonchelos que eran sobrios en Viena.  
Roberto dijo que compró unos violonchelos que eran sobrio en Viena.  
Roberto dijo que compró unos violonchelos que eran sobrias en Viena.
86. Lucía dijo que encontró un piano que era viejo en Austria.  
Lucía dijo que encontró un piano que era vieja en Austria.  
Lucía dijo que encontró un piano que era viejos en Austria.  
Lucía dijo que encontró unos pianos que eran viejos en Austria.  
Lucía dijo que encontró unos pianos que eran viejo en Austria.  
Lucía dijo que encontró unos pianos que eran viejas en Austria.
87. Lucía dijo que vio un mercado que era festivo en Andalucía.  
Lucía dijo que vio un mercado que era festiva en Andalucía.  
Lucía dijo que vio un mercado que era festivos en Andalucía.  
Lucía dijo que vio unos mercados que eran festivos en Andalucía.  
Lucía dijo que vio unos mercados que eran festivo en Andalucía.  
Lucía dijo que vio unos mercados que eran festivas en Andalucía.
88. Luis dijo que compró un diccionario que era bueno en México.  
Luis dijo que compró un diccionario que era buena en México.  
Luis dijo que compró un diccionario que era buenos en México.  
Luis dijo que compró unos diccionarios que eran buenos en México.  
Luis dijo que compró unos diccionarios que eran bueno en México.  
Luis dijo que compró unos diccionarios que eran buenas en México.
89. Luis dijo que comió un bocadillo que era sano en Rusia.  
Luis dijo que comió un bocadillo que era sana en Rusia.  
Luis dijo que comió un bocadillo que era sanos en Rusia.  
Luis dijo que comió unos bocadillos que eran sanos en Rusia.  
Luis dijo que comió unos bocadillos que eran sano en Rusia.  
Luis dijo que comió unos bocadillos que eran sanas en Rusia.
90. Natalia dijo que vio un concierto que era alternativo en California.  
Natalia dijo que vio un concierto que era alternativa en California.  
Natalia dijo que vio un concierto que era alternativos en California.  
Natalia dijo que vio unos conciertos que eran alternativos en California.  
Natalia dijo que vio unos conciertos que eran alternativo en California.  
Natalia dijo que vio unos conciertos que eran alternativas en California.

91. Natalia dijo que comió un plátano que era ácido en Escocia.  
Natalia dijo que comió un plátano que era ácida en Escocia.  
Natalia dijo que comió un plátano que era ácidos en Escocia.  
Natalia dijo que comió unos plátanos que eran ácidos en Escocia.  
Natalia dijo que comió unos plátanos que eran ácido en Escocia.  
Natalia dijo que comió unos plátanos que eran ácidas en Escocia.
92. Andrés dijo que comió un queso que era exótico en Francia.  
Andrés dijo que comió un queso que era exótica en Francia.  
Andrés dijo que comió un queso que era exóticos en Francia.  
Andrés dijo que comió unos quesos que eran exóticos en Francia.  
Andrés dijo que comió unos quesos que eran exótico en Francia.  
Andrés dijo que comió unos quesos que eran exóticas en Francia.
93. Andrés dijo que comió un huevo que era rosado en Sidney.  
Andrés dijo que comió un huevo que era rosada en Sidney.  
Andrés dijo que comió un huevo que era rosados en Sidney.  
Andrés dijo que comió unos huevos que eran rosados en Sidney.  
Andrés dijo que comió unos huevos que eran rosado en Sidney.  
Andrés dijo que comió unos huevos que eran rosadas en Sidney.
94. Arturo dijo que comió un espárrago que era graso en China.  
Arturo dijo que comió un espárrago que era grasa en China.  
Arturo dijo que comió un espárrago que era grasos en China.  
Arturo dijo que comió unos espárragos que eran grasos en China.  
Arturo dijo que comió unos espárragos que eran graso en China.  
Arturo dijo que comió unos espárragos que eran grasas en China.
95. Claudia dijo que encontró un contrato que era justo en Alemania.  
Claudia dijo que encontró un contrato que era justa en Alemania.  
Claudia dijo que encontró un contrato que era justos en Alemania.  
Claudia dijo que encontró unos contratos que eran justos en Alemania.  
Claudia dijo que encontró unos contratos que eran justo en Alemania.  
Claudia dijo que encontró unos contratos que eran justas en Alemania.
96. Claudia dijo que encontró un pasatiempo que era divertido en Tanzania.  
Claudia dijo que encontró un pasatiempo que era divertida en Tanzania.  
Claudia dijo que encontró un pasatiempo que era divertidos en Tanzania.  
Claudia dijo que encontró unos pasatiempos que eran divertidos en Tanzania.  
Claudia dijo que encontró unos pasatiempos que eran divertido en Tanzania.  
Claudia dijo que encontró unos pasatiempos que eran divertidas en Tanzania.

97. Nuria dijo que encontró un utensilio que era primitivo en Finlandia.  
Nuria dijo que encontró un utensilio que era primitiva en Finlandia.  
Nuria dijo que encontró un utensilio que era primitivos en Finlandia.  
Nuria dijo que encontró unos utensilios que eran primitivos en Finlandia.  
Nuria dijo que encontró unos utensilios que eran primitivo en Finlandia.  
Nuria dijo que encontró unos utensilios que eran primitivas en Finlandia.
98. Nuria dijo que vio un espectáculo que era violento en Sevilla.  
Nuria dijo que vio un espectáculo que era violenta en Sevilla.  
Nuria dijo que vio un espectáculo que era violentos en Sevilla.  
Nuria dijo que vio unos espectáculos que eran violentos en Sevilla.  
Nuria dijo que vio unos espectáculos que eran violento en Sevilla.  
Nuria dijo que vio unos espectáculos que eran violentas en Sevilla.
99. Diego dijo que vio un juego que era aburrido en Portugal.  
Diego dijo que vio un juego que era aburrida en Portugal.  
Diego dijo que vio un juego que era aburridos en Portugal.  
Diego dijo que vio unos juegos que eran aburridos en Portugal.  
Diego dijo que vio unos juegos que eran aburrido en Portugal.  
Diego dijo que vio unos juegos que eran aburridas en Portugal.
100. Diego dijo que visitó un edificio que era nuevo en Brasil.  
Diego dijo que visitó un edificio que era nueva en Brasil.  
Diego dijo que visitó un edificio que era nuevos en Brasil.  
Diego dijo que visitó unos edificios que eran nuevos en Brasil.  
Diego dijo que visitó unos edificios que eran nuevo en Brasil.  
Diego dijo que visitó unos edificios que eran nuevas en Brasil.
101. Rogelio dijo que descubrió un equipo que era agresivo en Alemania.  
Rogelio dijo que descubrió un equipo que era agresiva en Alemania.  
Rogelio dijo que descubrió un equipo que era agresivos en Alemania.  
Rogelio dijo que descubrió unos equipos que eran agresivos en Alemania.  
Rogelio dijo que descubrió unos equipos que eran agresivo en Alemania.  
Rogelio dijo que descubrió unos equipos que eran agresivas en Alemania.
102. Rogelio dijo que descubrió un vídeo que era formativo en Canadá.  
Rogelio dijo que descubrió un vídeo que era formativa en Canadá.  
Rogelio dijo que descubrió un vídeo que era formativos en Canadá.  
Rogelio dijo que descubrió unos vídeos que eran formativos en Canadá.  
Rogelio dijo que descubrió unos vídeos que eran formativo en Canadá.  
Rogelio dijo que descubrió unos vídeos que eran formativas en Canadá.

103. Carmela dijo que descubrió un armario que era hondo en Finlandia.  
Carmela dijo que descubrió un armario que era honda en Finlandia.  
Carmela dijo que descubrió un armario que era hondos en Finlandia.  
Carmela dijo que descubrió unos armarios que eran hondos en Finlandia.  
Carmela dijo que descubrió unos armarios que eran hondo en Finlandia.  
Carmela dijo que descubrió unos armarios que eran hondas en Finlandia.
104. Carmela dijo que descubrió un dibujo que era serio en Amsterdam.  
Carmela dijo que descubrió un dibujo que era seria en Amsterdam.  
Carmela dijo que descubrió un dibujo que era serios en Amsterdam.  
Carmela dijo que descubrió unos dibujos que eran serios en Amsterdam.  
Carmela dijo que descubrió unos dibujos que eran serio en Amsterdam.  
Carmela dijo que descubrió unos dibujos que eran serias en Amsterdam.
105. Evelyn dijo que descubrió un exorcismo que era destructivo en Habana.  
Evelyn dijo que descubrió un exorcismo que era destructiva en Habana.  
Evelyn dijo que descubrió un exorcismo que era destructivos en Habana.  
Evelyn dijo que descubrió unos exorcismos que eran destructivos en Habana.  
Evelyn dijo que descubrió unos exorcismos que eran destructivo en Habana.  
Evelyn dijo que descubrió unos exorcismos que eran destructivas en Habana.
106. Evelyn dijo que descubrió un gobierno que era autoritario en China.  
Evelyn dijo que descubrió un gobierno que era autoritaria en China.  
Evelyn dijo que descubrió un gobierno que era autoritarios en China.  
Evelyn dijo que descubrió unos gobiernos que eran autoritarios en China.  
Evelyn dijo que descubrió unos gobiernos que eran autoritario en China.  
Evelyn dijo que descubrió unos gobiernos que eran autoritarias en China.
107. Alejandro dijo que descubrió un ajo que era dañino en Japón.  
Alejandro dijo que descubrió un ajo que era dañina en Japón.  
Alejandro dijo que descubrió un ajo que era dañinos en Japón.  
Alejandro dijo que descubrió unos ajos que eran dañinos en Japón.  
Alejandro dijo que descubrió unos ajos que eran dañino en Japón.  
Alejandro dijo que descubrió unos ajos que eran dañinas en Japón.
108. Alejandro dijo que descubrió un matrimonio que era poderoso en Cuba.  
Alejandro dijo que descubrió un matrimonio que era poderosa en Cuba.  
Alejandro dijo que descubrió un matrimonio que era poderosos en Cuba.  
Alejandro dijo que descubrió unos matrimonios que eran poderosos en Cuba.  
Alejandro dijo que descubrió unos matrimonios que eran poderoso en Cuba.  
Alejandro dijo que descubrió unos matrimonios que eran poderosas en Cuba.

